

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

TŘÍDĚNÍ EKG CYKLŮ POMOCÍ SHLUKOVÉ ANALÝZY

CLUSTERING OF ECG CYCLES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Karolína Němečková

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Marina Ronzhina, Ph.D.

BRNO 2020

Bakalářská práce

bakalářský studijní obor **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

Studentka: Karolína Němečková

ID: 195162

Ročník: 3

Akademický rok: 2019/20

NÁZEV TÉMATU:

Třídění EKG cyklů pomocí shlukové analýzy

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši v oblasti projevů poruch srdečního rytmu v EKG záznamech. Dále se seznámte s metodami zpracování a analýzy EKG. Zaměřte se na parametry EKG (jak morfologické tak rytmické) a shlukovací algoritmy používané pro třídění cyklů EKG. 2) Seznamte se s databází patientských EKG záznamů dostupnou na ÚBMI. Popište patologie vyskytující se v záznamech z databáze. 3) Realizujte vybranou metodu třídění jednotlivých EKG cyklů v prostředí MATLAB. 4) Otestujte navrženou metodu na záznamech od pacientů s různými patologiemi. Vyhodnoťte úspěšnost použité metody. 5) Na základě předchozího bodu proveďte optimalizaci metody a automatizujte nastavení parametrů metody. 6) Dosažené výsledky vhodně interpretujte. Diskutujte výhody a limity navržené třídící metody.

DOPORUČENÁ LITERATURA:

[1] CURTIS, Michael J., Jules C. HANCOX, András FARKAS, et al. The Lambeth Conventions (II): Guidelines for the study of animal and human ventricular and supraventricular arrhythmias. Pharmacology and Therapeutics, Amsterdam: Elsevier Inc., 2013. roč. 139, vyd. 2, s. 213 – 248. ISSN 0163-7258.

[2] SMISEK, Radovan, Jakub HEJC, Marina RONZHINA, et al. Multi-stage SVM approach for cardiac arrhythmias detection in short single-lead ECG recorded by a wearable device. Physiological Measurements, 2018.

[3] YEH, Yun-Chi, Che W. CHIOU, Hong-Jhih LIN. Analyzing ECG for cardiac arrhythmia using cluster analysis. Expert Systems with Applications, 2012. roč. 39, vyd. 1, s. 1000-1010.

Termín zadání: 3.2.2020

Termín odevzdání: 5.6.2020

Vedoucí práce: Ing. Marina Ronzhina, Ph.D.

prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTRAKT

Tato bakalářská práce se zabývá aplikací shlukové analýzy na různé záznamy EKG s cílem rozpoznat jednotlivé patologie se zaměřením na detekci komorových a síňových extrasystol. Práce je založena na korelaci signálů a následné identifikaci patologií pomocí základních charakteristik jednotlivých rytmů.

Metoda byla vyhodnocena na testovacích datech a dosáhla TPR 73.40 %, FPR 91.00 %, PPV 29.00 %, ACC 90.00 %, F1 41.40 % pro detekci PAC a TPR 76.50 %, FPR 94.20 %, PPV 45.90 %, ACC 93.10 %, F1 57.40 % pro detekci PVC. Nízké hodnoty F1 a PPV jsou způsobeny vysokým počtem falešně pozitivních detekcí vyskytujících se hlavně v zašuměných EKG záznamech a záznamech s přítomností atriální fibrilace.

KLÍČOVÁ SLOVA

EKG, RR interval, korelace, detekce extrasystol, shlukování srdečních rytmů, fuzzy inferenční systém

ABSTRACT

This bachelor thesis deals with application of cluster analysis to different ECG records in order to identify particular cardiac pathologies. The work is mainly focused on the detection of premature atrial and premature ventricular beats. Presented approach is based on the signal correlation and further beat type identification and beats clustering via specific ECG features.

By evaluation the method on test data, we obtained TPR 73.40 %, FPR 91.00 %, PPV 29.00 %, ACC 90.00 %, F1 41.40 % for PAC detection and TPR 76.50 %, FPR 94.20 %, PPV 45.90 %, ACC 93.10 %, F1 57.40 % for PVC detection. Pure F1 and PPV is due to high number of false positive detections mainly in noisy ECG or ECG with manifested atrial fibrillation.

KEYWORDS

ECG, RR interval, correlation, extrasystols detection, cardiac beats clustering, fuzzy inference system

NĚMEČKOVÁ, Karolína. *Třídění EKG cyklů pomocí shlukové analýzy*. Brno, 2020, 62 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce: Ing. Marina Ronzhina, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Třídění EKG cyklů pomocí shlukové analýzy“ jsem vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autorka uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autorky

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu bakalářské práce Ing. Marině Rozhině, Ph.D. za odborné vedení, konzultace, trpělivost, podnětné návrhy k práci a cenné rady.

Obsah

Úvod	10
1 Elektrocardiogram	12
2 Moderní metody zpracování záznamu EKG	18
3 Shluková analýza dat	21
4 Rozhodování s využitím fuzzy logiky	25
5 Popis analyzovaných dat	27
6 Programové řešení	30
6.1 Primární filtrace dat	30
6.2 Optimalizace vrcholu R kmitu	31
6.3 Segmentace signálu	32
6.4 Shlukování na základě korelace	34
6.5 Detekce síňových extrasystol	37
6.6 Výběr příznaků	39
6.7 Detekce komorových extrasystol na základě fuzzy usuzování	42
7 Vyhodnocení algoritmu	47
Závěr	55
Literatura	56
Seznam symbolů, veličin a zkratk	59
A ROC křivky a boxploty vybraných příznaků	60

Seznam obrázků

1.1	Izolovaná síňová extrasystola	13
1.2	Komorová extrasystola	14
1.3	Atrioventrikulární blokáda II.stupně	15
1.4	Atrioventrikulární blokáda III. stupně	16
1.5	Blokáda levého Tawarova raménka	16
1.6	Fibrilace síní	17
3.1	Obecné schéma shlukové analýzy	21
4.1	Ukázka různých funkcí příslušnosti	25
5.1	Ukázka zašuměného signálu	28
5.2	Nepřesná detekce vrcholů R kmitů	28
5.3	Nedetekované QRS komplexy	29
5.4	Detekce QRS komplexů i v místech, kde se nenachází	29
6.1	Obecné schéma programového řešení	30
6.2	Optimalizace vrcholu R	32
6.3	Segmentace záznamu	33
6.4	Doplnění segmentu na začátku záznamu	34
6.5	Přehledové schéma shlukování	35
6.6	Ukázka shlukování na základě korelace u minimálně zašuměných dat	37
6.7	Ukázka shlukování na základě korelace u vysoce zašuměných dat	37
6.8	Správná detekce PAC	39
6.9	Problém detekce PAC u atrio-ventrikulární disociace rytmů	39
6.10	Ukázka zlepšení příznaku	40
6.11	Ukázka získání příznaku	43
6.12	Funkce příslušnosti	44
6.13	Výstupní funkce příslušnosti	45
6.14	Ukázka správné funkce programu pro detekci PVC	46
7.1	Příčiny falešně pozitivních detekcí PAC	49
7.2	Příčiny falešně pozitivních detekcí PVC	50
7.3	Falešně pozitivní detekce PAC způsobená záměnou PVC za PAC	51
7.4	Falešně pozitivní detekce PAC	52
7.5	Falešně pozitivní detekce PAC způsobená nesprávnou detekcí R kmitu	52
7.6	Ukázka chybné funkce shlukování	53
7.7	Ukázka správné funkce klasifikátoru	53
7.8	Ukázka správné funkce klasifikátoru (2)	54
A.1	Příznak 1	60
A.2	Příznak 2	60
A.3	Příznak 3	61

A.4	Příznak 4	61
A.5	Příznak 5	62
A.6	Příznak 7	62

Seznam tabulek

1.1	Přehled standardních hodnot EKG záznamu	13
5.1	Rozdělení signálů do jednotlivých kategorií	27
7.1	Tabulka PAC a PVC pro trénovací data	47
7.2	Tabulka PAC a PVC pro testovací data	47
7.3	Výsledky detekce PAC	48
7.4	Výsledky detekce PVC	48
7.5	Výsledky detekce extrasystol obecně	48

Úvod

V kardiologické praxi patří elektrokardiografie mezi základní vyšetřovací metody. Důvodem je nejenom naprostá neinvazivnost zmiňované techniky, ale i informační výtěžnost provedeného vyšetření. Každý pacient může mít tvar EKG křivky drobně přizpůsobený své postavě (výška postavy ovlivňuje šířku QRS komplexu, tloušťka výšku kmitů), avšak základní charakteristiky fyziologického i patologického EKG bývají většinou zachovány a proto je pro tuto metodu automatické strojové zpracování vhodnou volbou.

Ektopický stah či extrasystola jsou z praktického pohledu synonyma pro specifickou srdeční aritmii. Tato aritmie je definována jako stah, který přichází dříve, než v řadě další očekávaný stah. Místo vzniku takového stahu může být dvojí, buď vzniká v srdečních síních a jedná se o supraventrikulární (síňovou) extrasystolu, nebo jsou místem vzniku srdeční komory a jedná se o komorovou extrasystolu. Oba z výše zmíněných stahů mají svá specifika, podle kterých mohou být po pečlivém prozkoumání záznamu EKG rozpoznány. Co se týče závažnosti těchto aritmií, pak extrasystoly můžeme pozorovat jak u zdravých jedinců tak i u lidí se závažným srdečním onemocněním, kde jsou extrasystoly (převážně komorové) spojeny se zvýšenou kardiiovaskulární mortalitou.

Cílem práce je navrhnout a realizovat algoritmus na základě metod shlukové analýzy, který bude provádět třídění jednotlivých EKG cyklů. Práce se zaměřuje zvláště na síňové a komorové extrasystoly, které by měly být pomocí algoritmu nalezeny a přesně lokalizovány. Detekce extrasystol by mohla přispět k časnému odhalení závažných postižení srdce, jako je například fibrilace komor, jejichž indikátorem mohou být právě extrasystoly. Většina dosavadních vyhodnocovacích systémů pracují na základech umělé inteligence a strojového učení, což sice vede k dosažení kvalitních výsledků detekce, ale za cenu složité či zcela nemožné interpretace rozhodovacího procesu a tím i výsledků. Dílčím cílem této práce je co nejvíce přiblížit rozhodování tohoto systému postupu zdravotnického personálu při vyhodnocování záznamu EKG. Předpokládaným výsledkem je číselný vektor, kde jsou QRS komplexy zastoupeny čísly, která označují shluk, do něhož byl cyklus zařazen. Cykly nezařazené mezi extrasystoly (komorové či síňové) spadají do shluku třetího, který již nepředstavuje žádnou konkrétní patologii.

Celá práce se zakládá na hlavní charakteristice extrasystol. Předčasné síňové stahy (premature atrial contraction, PAC) se v EKG projevují změnou morfologie P vlny bez změny QRS komplexu či vlny T. Dále také odpovídající RR intervaly v okamžiku výskytu extrasystoly mají charakteristické vlastnosti, konkrétně krátký extrasystolický a dlouhý postextrasystolický RR interval. Komorové extrasystoly (premature ventricular contraction, PVC) se liší morfologií QRS komplexu, absencí P vlny

a odlišnou délkou RR intervalů. Charakteristika vyplývá ze samotného vzniku a šíření elektrického impulzu, který je generován v komoře namísto sinusového uzlu, jak tomu bývá za normálních okolností.

Začátek práce je věnován popisu vzniku a šíření srdečního vzruchu a jednotlivým patologiím, které jsou důležité pro vybranou problematiku. Dále je pro získání obecného přehledu do práce zahrnuta kapitola shrnující moderní metody detekce extrasystol v EKG. V další kapitole jsou stručně představeny metody shlukové analýzy se zaměřením na metodu k-means, která je v této oblasti standardní metodou. Kapitola pokračuje popisem shlukovací metody na základě korelace, která byla pro tuto práci vybrána.

Následující kapitoly jsou již věnovány realizaci algoritmu v programovém prostředí MATLAB, který je pro tyto účely vhodný. Začátek se věnuje popisu použitých dat a následně se přechází na jednotlivé části algoritmu, řešící dílčí problémy. Dosažené výsledky jsou diskutovány v kapitole 7. Práce je zakončena závěrem, který obsahuje pojednání o splnění cíle a jednotlivých předpokladů.

1 Elektrokardiogram

Základní vlastností srdečních buněk je jejich elektrická aktivita, kterou můžeme vyjádřit pomocí EKG křivky. Tvar elektrokardiogramu má vysokou vypovídající schopnost. Skládá se z vln a kmitů, které na sebe navazují a každá část odpovídá jiné fázi srdečního cyklu. Podle pořadí hovoříme o vlně P, komplexu QRS (skládající se ze tří kmitů), vlně T a U. Podle místa změny křivky pak můžeme odvodit nejen druh srdeční poruchy, ale i její lokaci. V rámci hodnocení se zkoumají morfologické změny jednotlivých vln a kmitů spolu s jejich časovými intervaly. Mezi často měřené patří intervaly PR a PQ, komplex QRS a úsek ST.

Zdrojem obsahu kapitoly jsou [1][2][3].

Fyziologické EKG

V EKG u zdravého jedince vzruch vzniká v SA (sinoatriálním) uzlu a šíří se internodálními síňovými spoji přes AV (atrioventrikulární) uzel, který je jediným vodivým spojením mezi síněmi a komorami. Dále prochází přes Hisův svazek na pravé a levé Tawarovo raménko. Pravé Tawarovo raménko vede vzruch do myokardu pravé komory. Levé Tawarovo raménko se dělí na přední a zadní fasciculus, které odpovídají za vedení vzruchu k interventrikulárnímu septu a myokardu levé komory. Tato raménka se následně větví na Purkyňova vlákna a vedou vzruch na pracovní myokard. Výše popsaná cesta šíření vzruchu je nejrychlejší možná a nazývá se převodní systém srdeční.

Jak již bylo zmíněno dříve elektrokardiogram se skládá z několika základních vln a kmitů. Vlna P odpovídá depolarizaci síní, dále vzruch pokračuje převodním systémem srdečním do komor. Depolarizaci komor odpovídá komplex QRS, za kterým je také skryta repolarizace síní. Vlna T pak představuje komorovou repolarizaci. Význam vlny U zatím není zcela znám.

Samotný tvar křivky může ovlivnit nejen vzniklá patologie, ale i stavba těla pacienta. I přes tento fakt se však v rámci EKG u zdravého pacienta pohybujeme v jistých standardních hodnotách (viz tabulka 1.1).

Patologické EKG

Pokud se EKG křivka liší od zavedeného standardu v intervalové či morfologické oblasti, může to značit jistou patologii srdeční aktivity. Porucha srdečního rytmu může mít velice variabilní příčiny. Podle místa vzniku dělíme poruchy srdečního rytmu na sinusové, síňové, junkční (v oblasti síňokomorového spojení) a komorové. Mohou se dělit také na základě mechanismu vzniku na poruchy tvorby vzruchu a poruchy vedení vzruchu.

Tab. 1.1: Přehled standardních hodnot EKG záznamu zdravého jedince [5].

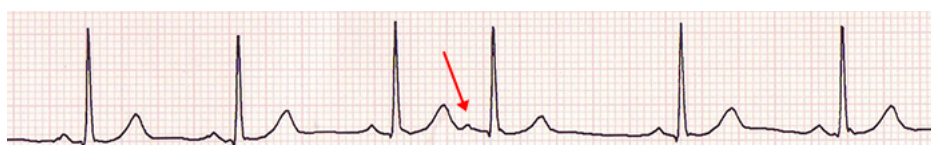
Interval	Normální hodnota	Rozptyl hodnot
P vlna	110 ms	± 20 ms
PQ interval	160 ms	± 40 ms
QRS komplex	100 ms	± 20 ms
QT interval	400 ms	± 40 ms
Amplituda vlny P	0,115 mV	$\pm 0,05$ mV
Amplituda komplexu QRS	1.5 mV	$\pm 0,5$ mV

Síňová extrasystola (PAC)

Síňová extrasystola je způsobena předčasnou depolarizací síně. Vzruch vzniká mimo oblast SA uzlu. Příčinou této předčasné excitace je zvýšená dráždivost síňových struktur. Vzhledem k místu vzniku vzruchu se signál v oblasti síní šíří atypicky, mimo převodní systém srdeční, následkem čehož dochází k aberaci tvaru vlny P na elektrokardiogramu. Úroveň aberace závisí na vzdálenosti ektopického střediska od SA uzlu. Vlna P se objevuje dříve, interval P-Q může být prodloužený a interval RR se zkracuje, jak je vidět na Obrázku 1.1. Při velice brzké excitaci může být pozorována superpozice vln T a P. V průběhu šíření vzruchu depolarizační vlna přechází i přes SA uzel, který tím vybije a nový vzruch se v SA uzlu tvoří de novo a nastává tzv. neúplná kompenzační pauza.

Přes AV uzel a zbývající část myokardu se vzruch šíří standardní cestou převodního systému srdečního. Komplex QRS je proto normální a kromě ojedinělých extrasystol zůstává SA uzel základním pacemakerem.

Síňové extrasystoly se nemusí objevovat pouze ojediněle, ale i v podobě tzv. kupletů či tripletů, což znamená více extrasystol následujících bezprostředně po sobě bez náznaku pravidelnosti. Extrasystoly se projevují i ve formě bigeminie, trigeminie popřípadě quadruminie, u kterých můžeme pozorovat výskyt předčasné depolarizace síní s danou pravidelností.



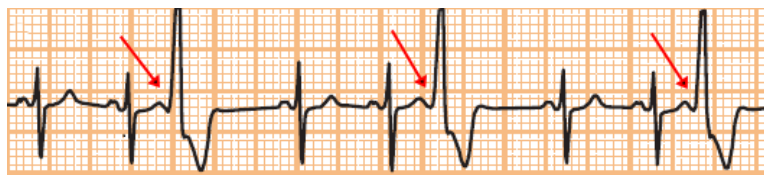
Obr. 1.1: Izolovaná síňová extrasystola [4].

Komorová extrasystola (PVC)

Předčasné komorové excitace vznikají distálně od bifurkace Hisova svazku a přicházejí dříve než další očekávaný sinusový vztah. Šíří se atypickou cestou přes pracovní myokard, což má za následek jejich široký, déle trvající komplex QRS. S rostoucí vzdáleností ektopického centra od AV uzlu se mění i morfologie QRS.

Vzhledem k lokaci vzniku excitace vlna P obvykle zcela chybí. Ojediněle se vzruch retrogradně převede přes AV uzel a vlna P vznikne, avšak na elektrokardiogramu není patrná, protože splývá s rozšířeným komplexem QRS. Tato předčasná excitace zpravidla neporuší základní sinusový rytmus, avšak ovlivní jeho následné šíření. Ve chvíli, kdy pravidelný sinusový vztah přijde, zastihne komory v refrakterní fázi a dál se nešíří. Komory jsou depolarizovány až v následujícím intervalu excitace SA uzlu. Tento děj má za následek vznik tzv. úplné postextrasystolické kompenzační pauzy. Ojediněle se může objevit komorová extrasystola vmezeřená mezi dva fyziologické QRS komplexy, avšak tento děj nastává pouze výjimečně u pomalého sinusového rytmu.

Komorové extrasystoly se mohou objevovat ojediněle, nebo jak již bylo zmíněno dříve, v podobě extrasystolické bigeminie, trigeminie (viz Obrázek 1.2), kupletu popřípadě tripletu.



Obr. 1.2: Příklad komorové trigeminie - každá třetí systola je komorová extrasystola s typickým rozšířeným komplexem QRS [4].

Atrioventrikulární blokáda (AVB)

Atrioventrikulární blokáda je způsobena poruchou převodního systému srdečního v oblasti AV uzlu či Hisova svazku. Nastává porucha vedení vzruchu mezi síněmi a komorami. Excitační vzruch vzniká v SA uzlu, ale převod na komory probíhá s jistou nepravidelností. Podle stupně závažnosti rozeznáváme atrioventrikulární blokádu I., II. a III.. stupně.

Atrioventrikulární blokáda I. stupně

Při atrioventrikulární bloádě I. stupně dochází ke zpomalenému vedení vzruchu přes AV uzel. Impulz přijde standardní cestou ze sinusového pacemakeru, ale při

převodu na komory se zdrží, následkem čehož dochází k prodloužení PQ intervalu na elektrokardiogramu. Tento stupeň blokády by během klasifikace extrasystol neměl být problémový.

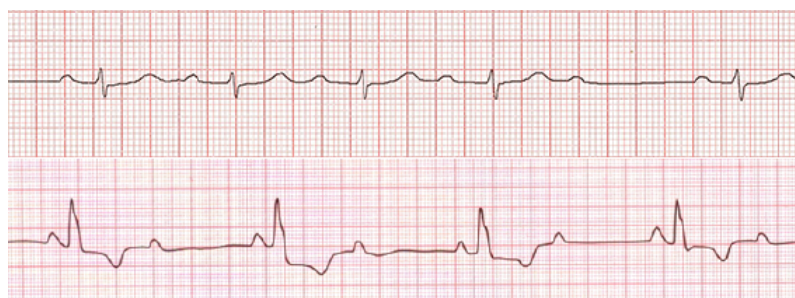
Atrioventrikulární blokáda II. stupně

Pokud dochází k úplnému přerušení převodu každého n-tého vzruchu, pak mluvíme o AVB II. stupně, kterou můžeme rozdělit na dva typy, AVB typu Mobitz a AVB typu Wenckebach, jejichž rozdíly jsou patrné na Obrázku 1.3.

Pokud jsou PR intervaly stále konstantní, ale najednou se signál nepřevede ze síní na komory mluvíme o blokádě typu Mobitz a tímto vzniklá pauza je rovna dvojnásobku klasického RR intervalu. Pokud se blokáda nachází za bifurkací Hisova svazku můžeme pozorovat rozšířený komplex QRS.

Naprosto odlišné projevy má blokáda typu Wenckebach u které pozorujeme postupné prodlužování PR intervalu a následné nepřevedení vzruchu. Poté se cyklus resetuje a začíná od znova.

Výše zmíněné stupně blokády mohou být problematické jak při klasifikaci síňové extrasystoly (nepravidelné RR intervaly), tak komorové extrasystoly (rozšířený komplex QRS při blokaci Hisova svazku za bifurkací).

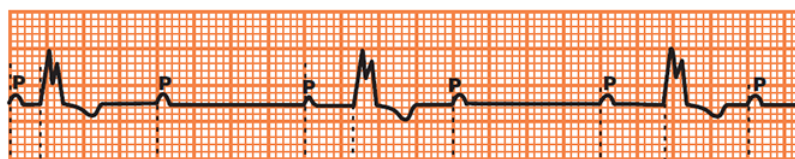


Obr. 1.3: AVB II.stupně - AVB typu Wenckebach s postupným prodlužováním PR intervalu a následným výpadkem (nahore), AVB typu Mobitz se širokým komplexem QRS způsobeným blokací za bifurkací Hisova svazku (dole) [4].

Atrioventrikulární blokáda III. stupně

Atrioventrikulární blokáda III. stupně je již kompletní blokáda AV uzlu, kdy jsou komory a síně elektricky izolované a na komory se nepřevádí žádný signál. Dochází k síňokomorové disociaci. Řešením této situace je spuštění náhradního srdečního rytmu pomocí ektopického ložiska pod blokádou. Ten může vycházet z AV uzlu za blokádou (junkční rytmus), který bude mít úzký komplex QRS. Pokud je však

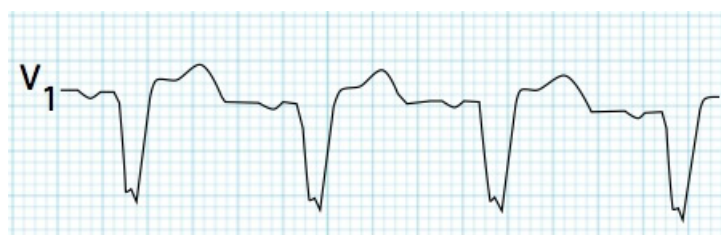
blokáda umístěna níže, dochází ke spuštění komorového rytmu, který se šíří pomaleji a projevuje se rozšířeným komplexem QRS (viz Obrázek 1.4). AVB III. stupně může být opět problematická z důvodu širokého komplexu QRS.



Obr. 1.4: AVB III. stupně - můžeme vidět široký komplex QRS značící komorový rytmus [4].

Blokády pravého a levého Tawarova raménka (BBB)

V případě blokády Tawarova raménka se signál šíří přes síně až do oblasti mezikomorového septa standardní cestou. Délky intervalů PR jsou tedy normální. Šíření vzruchu v komorách však probíhá z části přes pracovní myokard z důvodu blokace jednoho raménka. Např. v případě blokády levého Tawarova raménka se signál šíří bez zpoždění do pravé komory a poté přechází do zatím nedepolarizované levé komory pomocí pracovního myokardu (signál se do levé komory převodním systémem srdečním nedostal). Toto šíření je pomalejší a na záznamu se projevuje rozšířeným komplexem QRS (viz Obrázek 1.5), který, jak již bylo zmíněno dříve, by mohl snižovat procentuální úspěšnost klasifikace PVC. Podobný průběh můžeme vidět i v případě blokády pravého Tawarova raménka.



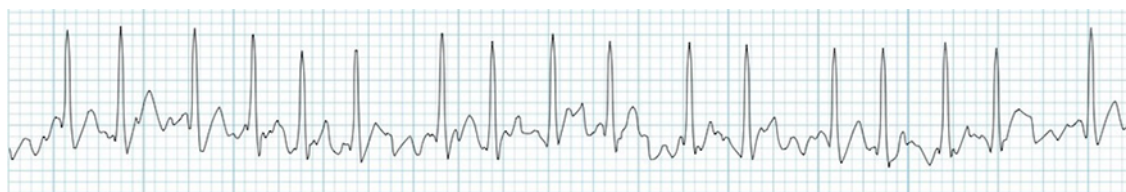
Obr. 1.5: Blokáda levého Tawarova raménka - můžeme vidět typický široký komplex QRS [4].

Supraventrikulární arytmie

Poslední důležitou oblastí patologií jsou pro tuto práci supraventrikulární arytmie. Jedná se o poruchy srdečního rytmu, které vznikají nad Hisovým svazkem. Místem vzniku arytmii tedy může být SA uzel, myokard síní nebo AV uzel. Všechny

supraventrikulární arytmie mají stejný způsob šíření depolarizační vlny v rámci komor, kde se šíří klasickou cestou a tím pádem nedochází k tvarovým změnám komplexu QRS. Tyto abnormální rytmy můžeme také rozdělit na bradykardie, tachykardie, flutter a fibrilace. V rámci této práce tvoří zajímavou oblast pouze fibrilace síní.

Fibrilace síní je arytmií, kde dochází k vzniku nepravidelných kontrakcí komor, což má za následek nepravidelnost celého srdečního rytmu. Počátek nepravidelnosti je již v síních, kde jednotlivá svalová vlákna kontrahují nezávisle na sobě a AV uzel je neustále vystavován depolarizačním vlnám. AV uzel není schopen propustit všechny depolarizační vlny a proto prochází jenom některé, které dále depolarizují komory. Propustnost AV uzlu není pravidelná a proto je následný komorový rytmus nepravidelný. Typickým znakem fibrilace síní je absence vlny P (viz Obrázek 1.6).



Obr. 1.6: Fibrilace síní - patrná absence vlny P a nepravidelnost RR intervalů [4].

2 Moderní metody zpracování záznamu EKG

Pro získání přehledu v oblasti moderních metod shlukování byla provedena literární rešerše. V rámci zaměření na detekci extrasystol v EKG záznamu nebyla rešerše provedena pouze v oblasti metod shlukování, ale byla rozšířena o detekce srdečních arytmií (tedy i extrasystol) obecně.

Postupně budou v této kapitole popsány všechny kroky zpracování EKG záznamu a možné přístupy k jednotlivým krokům. Klasifikace EKG záznamu se obecně skládá z (1) předzpracování záznamu, (2) segmentace signálu, (3) výběru příznaků a (4) klasifikace. Přičemž každý krok má značný vliv na finální kvalitu detekce. V rámci celého procesu detekce EKG arytmií je obvykle diskutován také způsob získávání EKG záznamu, avšak tato problematika není náplní práce a proto do rešerše nebyla zahrnuta.

Předzpracování záznamu

Cílem předzpracování je snížit šum v EKG záznamu a zároveň minimalizovat tvarové změny způsobené filtrací. Nejrozšířenější přístup filtrace je pomocí rekurzivních filtrů s konečnou impulzní charakteristikou, které přinášejí dobré výsledky u známých rušivých frekvencích (např. 50 Hz). V případech, kdy není známa frekvenční oblast šumu však tyto filtry není možné použít a přechází se například na adaptivní filtraci, která je často založena na neuronových sítích pro získání lepších výsledků. Některé nejmodernější metody detekce srdečních arytmií dokonce fázi předzpracování záznamu vynechávají a pracují se surovým záznamem viz Escalona-Moran et al. [8] [9]

Segmentace signálu

Segmentace signálu zahrnuje detekci QRS komplexů (R kmitů) a následné rozdělení signálu na segmenty kolem detekovaných QRS komplexů. Délka segmentu je závislá na řešeném problému. Segment může obsahovat jen QRS komplex, nebo i jeho okolí (jako je např. vlna P nebo ST úsek), vše záleží na druhu informace, kterou chceme ze záznamu získat.

Detailnější rozměření (vlny P a T) nebývá v rámci moderních metod pravidlem. Přestože rozměření těchto částí záznamu EKG přináší více informací pro klasifikaci jsou tyto detekce spíše zdrojem chyb. Pro zvýšení úspěšnosti segmentace signálu jsou používány digitální filtry pro odstranění šumu a kolísání nulové izoliny, následované nelineární transformací signálu zlepšující detekci kmitu R za pomoci prahu, který se neustále adaptuje na příchozí signál. Tento algoritmus je používán díky jeho jednoduchosti a kvalitním výsledkům. Používají se však i sofistikovanější metody založené

na neuronových sítích, vlnkové transformaci, bance filtrů či genetických algoritmech. V mnohých pracích je však část detekce QRS komplexu vynechána, protože se často pracuje s již anotovanou databází MIT-BIH. [5]-[8]

EKG příznaky a jejich výběr

Dílčí segmenty EKG záznamu je vhodné místo posloupnosti naměřených vzorků reprezentovat menším počtem číselných příznaků. Příznaky nám reprezentují typické vlastnosti objektu, na jejichž základě se provádí klasifikace. V literatuře však můžeme najít i práce, ve kterých byly segmenty reprezentovány posloupností naměřených vzorků (viz dále).

V rámci extrakce příznaků se používají metody v časové, frekvenční nebo časově - frekvenční oblasti. Příznaky se také mohou dělit na příznaky zaměřené na popis morfologie EKG křivky (amplitudy, sklony křivek apod.) a příznaky popisující srdeční rytmus (RR interval, délka QRS komplexu apod.), které jsou v literatuře nejčastěji používané. Můžeme najít práce založené na extrakci příznaků pouze z RR intervalů, či z jejich normalizovaných podob. Kromě RR intervalů se používají i zbylé intervaly komplexu QRS (viz Tabulka 1.1), přičemž nejčastěji používaný interval bývá délka celého komplexu QRS.

Dalším voleným přístupem je použití samotných vzorků signálu jako příznaků, ať už jejich nezredukované formy nebo použití například interpolace pro omezení počtu vzorků a tím pádem i příznaků. Často volenou metodou je i redukce příznaků pomocí PCA (angl. Principal Component Analysis, analýza hlavních komponent). Také využití vlnkové transformace přináší slibné výsledky díky získávání příznaků z časové i frekvenční oblasti a proto je při extrakci příznaků volena také. Některé výzkumy uvádí, že vlnková transformace je pro získávání příznaků nejlepší metodou.

Vektorkardiogram je odlišná forma záznamu elektrických projevů srdce, která spojuje informace ze dvou svodů do jedné dvoudimenzionální reprezentace. Pro detekci extrasystol je doporučována kombinace svodů II a V1. Věřící se, že vektorkardiogram je vhodnou volbou při detekci extrasystol. [5] [8]

Dalším moderním přístupem je využití tzv. hlubokých neuronových sítí. V tomto případě se extrakce příznaků provádí v rámci klasifikace, kdy na vstup můžeme vkládat surové či nějakým způsobem zpracované záznamy EKG a na výstupu získáváme zařazení do určité kategorie. Příznaky jsou extrahovány automaticky v rámci klasifikace, během které se signál postupně zpracovává v jednotlivých vrstvách neuronové sítě.

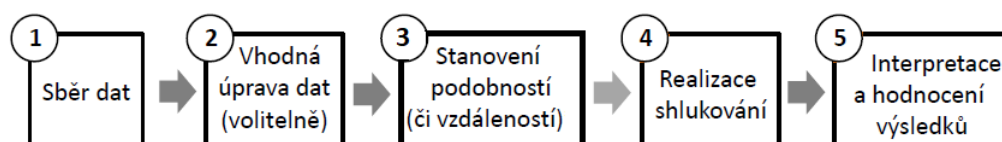
Klasifikace

Klasifikace je prováděna formou učících algoritmů. Nejpoužívanějšími algoritmy, pokud se bavíme o metodách využívající extrakci příznaků, jsou: metoda podpůrných vektorů (angl. support vector machine, SVM), umělé neuronové sítě (angl. artificial neural networks, ANN), lineární diskriminant a výpočet pomocí logistické regrese (angl. Reservoir Computing With Logistic Regression, RC). Avšak nejmodernější metody využívané pro klasifikaci jsou v dnešní době založeny na hlubokých neuronových sítích. Jak již bylo zmíněno, v těchto případech vynecháváme extrakci příznaků, protože ta je řešena v rámci jednotlivých vrstev. [8] [10]

Samozřejmě je zde možnost přístupů odlišných od metod umělé inteligence a výše zmíněných algoritmů. Klasifikaci je možné provádět i na základě výsledků shlukování. V tomto případě vycházíme z předpokladu, že podobné signály spadají do stejných shluků a dále provádět klasifikaci pouze v rámci shluků. Podobný přístup je volen v této práci.

3 Shluková analýza dat

Úkolem shlukové analýzy je seskupování objektů na základě jejich podobnosti. Základním pravidlem shlukování je, aby si objekty ve stejném shluku, byly více podobné, než objekty ze shluků různých. Obecné schéma shlukové analýzy můžeme vidět na obrázku 3.1. [12]



Obr. 3.1: Obecné schéma shlukové analýzy [11].

Prvním krokem shlukové analýzy je sběr dat, z kterých se následně získávají příznaky, podle kterých proběhne rozřazení do jednotlivých shluků. Na výběr příznaků se musí klást velký důraz, protože chybně zvolené příznaky mohou vést k nedostatečným výsledkům samotného shlukování. [12]

Po výběru příznaků následuje jejich úprava, která není nezbytná, avšak kvalitní úprava vede k hodnotnějším výsledkům. Jestliže jsou značné rozdíly mezi hodnotami jednotlivých příznaků, pak je vhodné upravit data pomocí standardizace, díky které se zbavíme nerovnoměrného vlivu příznaků na výpočet míry vzdálenosti, či podobnosti, podle námi zvolené metricky.

Základní vzdálenostní metrikou je výpočet Euklidovské vzdálenosti, avšak pokud máme i přes standardizaci vysoce variabilní data, je vhodné přejít k volbě jiné metricky. [11]

Následuje samotná realizace shlukování, která se liší dle použité metody.

Metody shlukové analýzy

Sběr a úprava dat je společná pro všechny metody shlukové analýzy. Základní rozdíl v jednotlivých metodách spočívá ve volbě objektů, mezi kterými se vzdálenost stanovuje, a obecným přístupem při tvorbě shluků. Většinou jsou rozeznávány dvě základní skupiny metod, kterými jsou metody hierarchické a metody nehierarchické. V rámci těchto metod probíhá rozdělení do jednotlivých shluků na základě binarity, prvek do shluku patří, nebo nepatří (1/0). Odlišný pohled na shlukování přináší tzv. fuzzy (měkké) shlukovací metody, které zařazují každý objekt do všech možných shluků, avšak s odlišným stupněm příslušnosti. Tyto metody fuzzy shlukování nebudou dále rozebírány. [11]

Hierarchické metody shlukování

Hierarchické metody jsou založeny na principu postupného seskupování prvků, přičemž se počítá s přirozeně strukturovanými daty, které mají dobře oddělitelné shluky. Jen v takovém případě je vhodné zvolit hierarchické shlukovací metody. Výsledky shlukování bývají nejčastěji vyjádřeny dendrogramy, což jsou vývojové stromy umožňující snadnou orientaci v proběhlém shlukování.

Princip shlukování spočívá ve výpočtu distanční matice z příznaků každého objektu, informující o vzdálenosti mezi jednotlivými objekty. Nejčastěji se jedná o matici Euklidovských vzdáleností. Z této matice se vybere minimum představující dva nejbližší objekty, které se následně spojí do jednoho shluku. Dalším krokem je přepočítání distanční matice, kde bude vytvořený shluk figurovat jako nový objekt. Právě v tomto kroku se hierarchické metody navzájem liší a to má za následek více možných alternativ shlukování.

Algoritmus probíhá do té doby, dokud nedojde ke spojení posledních dvou shluků do jednoho. Finálním krokem může být výpočet Pearsonova korelačního koeficientu (viz kapitola 3) pro zhodnocení zkreslení způsobeného dendrogramem. [11]

Nehierarchické metody shlukování

Pokud mají data hierarchickou strukturu, tvoří přirozeně odlišné shluky a nejedná se o velké množství dat jsou hierarchické metody vhodnou volbou pro shlukování. V opačném případě řešení realizujeme pomocí metod nehierarchických, které data nejčastěji rozkládají na předem známý počet shluků (není to však pravidlem). Nehierarchické metody jsou iterační optimalizační metody, kde v každém kroku iterace probíhá přerazení objektů do shluků. Počet iterací může být řízen různými přístupy, jako je například sledováním tzv. kritériální funkce informující nás o kvalitě rozřazení do shluků například pomocí součtu druhých mocnin vzdáleností všech objektů od jejich shluku. Kritériální funkce slouží ke sledování kvality shlukování a také k ukončení iteračního procesu, kdy po dosažení dostatečně nízké hodnoty kritériální funkce proces ukončujeme. Díky sledování kritériální funkce se postupně dostáváme k optimálnímu rozložení do shluků u kterých platí, že dva objekty ze stejného shluku si jsou blíže než dva objekty ze shluků různých. [11]

Standardní metodou nehierarchických shlukovacích metod je metoda **k-means** (metoda k-středů). Počet shluků k je v této metodě vždy předem znám a shluky jsou tvořeny kolem tzv. centroidů, které představují jistý reprezentativní objekt daného shluku. Centroid tedy představuje geometrický střed shluku, který vypočteme jako průměr objektů v rámci jednotlivých shluků. Počáteční umístění centroidů je zcela náhodné a iterativním postupem se jejich umístění optimalizuje tak, aby se zvýšila podobnost objektů v rámci shluků a zvýšila odlišnost objektů z různých

shluků. Metrikou podobnosti (v tomto případě vzdálenosti) je Euklidovská vzdálenost mezi objektem a centroidem. V průběhu jedné iterace proběhne výpočet vzdálenosti mezi každým objektem a každým centroidem a následně se objekt přiřadí ke shluku, k jehož centroidu je nejbližší. Protože přiřazením objektů se změnily geometrické středy shluků proběhne jejich přepočítání. Iterace probíhá dokud se rozložení do shluků mění, nebo není dosažena požadovaná hodnota kritériální funkce. Výrazným nedostatkem této metody je potřeba nastavení počtu shluků před provedením shlukování, a to i v případě, kdy nevíme do kolika shluků by bylo ideální shlukování provést. V případě nastavení nevhodného počtu mohou být výsledky naprosto nesmyslné. Dalším nedostatkem metody může být závislost výsledku shlukování na původním rozložení centroidů. [11] [12]

Shlukování na základě korelace

I přes nesporné výhody metody k-means, která je standardní metodou pro shlukování, jsem tuto metodu nezvolila. Pro odlišnou metodu jsem se rozhodla z důvodu značných nevýhod metody k-means. U této metody je potřeba hledat správnou hodnotu počtu shluků, kterou však dopředu neznáme a algoritmy pro hledání ideálního počtu shluků mohou u velice nízkého počtu objektů, s kterými zde pracuji, selhávat. Tato metoda shlukování je vhodná pro velké soubory dat, ve kterých jsou data rovnoměrně rozloženy do shluků a v opačných případech opět může docházet k získávání chybných výsledků.

Celkový výsledek je také ovlivněn počátečním nastavením centroidů a jejich pořadím. Pokud je objekt ve stejné vzdálenosti od dvou centroidů, zvolí první v pořadí jako bližší. Vzdálenost od centroidů se nejčastěji určuje na základě výpočtu míry vzdálenosti, přičemž u značně variabilních dat se na výsledku podílí převážné příznaky s větším rozdílem mezi objekty a dostáváme zkreslené výsledky, nevypovídající o skutečnosti. [11]

Z výše zmíněných důvodů nejdu touto cestou použití standardních metod a zkouším vyvinout vlastní metodu o které usuzuji, že by mohla dávat stejně kvalitní, či dokonce lepší výsledky a která by byla více přiblížena lidskému usuzování v praxi.

Stěžejním krokem navrženého algoritmu je shlukování na základě korelace dvou signálů pro popis jejich podobnosti, kdy nejprve bude využita vzájemná korelační funkce pro zjištění posunu, při kterém si jsou signály nejvíce podobné a následně bude proveden výpočet Piersonova korelačního koeficientu (právě pro vybraný posun) pro samotnou kvantifikaci podobnosti. Z tohoto důvodu je zde uvedena tato kapitola stručně popisující teoretický základ výpočtů. Předpoklady pro níže popsané postupy jsou v čase neměnné statistické vlastnosti signálu (tzv. stacionarita) a skutečnost, že tyto statistické vlastnosti mohou být zjištěny z jedné dostatečně dlouhé

realizace (tzv. ergodicita), které v tomto případě považujeme za splněné a můžeme hovořit o stacionárním ergodickém náhodném signálu. [13]

Vzájemná korelační funkce

Vzájemná korelace se snaží o kvantitativní vyjádření vazby mezi hodnotami signálu při určitém časovém posunu jednotlivých signálů. Při výpočtu korelační funkce se počítají korelační koeficienty pro všechny posuny. Výpočet je realizován pomocí vzorce

$$R_{xy}(\tau) = \frac{1}{M} \sum_{n=1}^{N-\tau} x(n)y(n+\tau) \quad (3.1)$$

kde se jedná o sumu součinů vzorků jednotlivých signálů (x, y) vždy s definovaným posunem (τ) , podělenou jejich počtem (M) . V tomto případě byl uveden vzorec pro výpočet korelační funkce, kde nebereme v potaz náběh a ukončení korelace, kde nedochází k překryvu všech vzorků. Délka signálu reprezentována písmenem M zůstává stále stejná a to má za následek nízké hodnoty amplitud signálu na okrajích korelační funkce. Pokud chceme brát v potaz částečné překrývání signálů používá se výpočet, kde se jmenovatel stále mění dle aktuálního překryvu. Vzorec je pak upraven a vypadá následovně

$$R_{xy}(\tau) = \frac{1}{M-\tau} \sum_{n=1}^{N-\tau} x(n)y(n+\tau) \quad (3.2)$$

avšak i tento přístup má své nevýhody, kterými jsou vyšší hodnoty šumu na okrajích signálu. [11][13]

Pearsonův korelační koeficient

Pearsonův korelační koeficient se nejčastěji používá pro měření lineární závislosti dvou náhodných veličin s konkrétním časovým posunem. Tento koeficient řeší problém kvantifikace hodnot u výpočtu korelace, kdy při výpočtu korelační funkce neznáme její meze a z toho důvodu nemá vysoké vypovídající schopnosti, pokud by jsme prováděli prahování podobnosti na základě hodnot získaných právě pomocí této korelace. Pearsonův korelační koeficient je normalizován a nabývá hodnot v rozmezí $<-1;1>$. Definice vypadá následovně

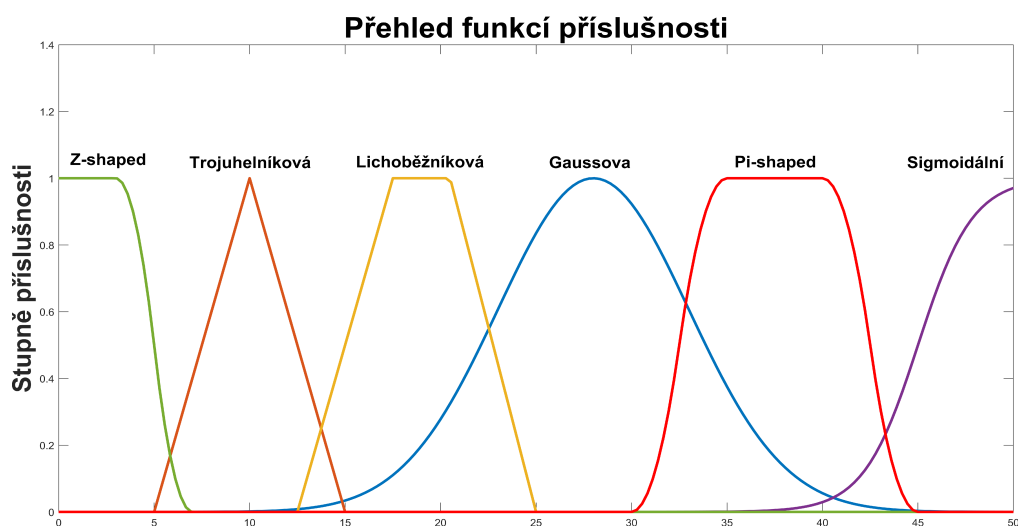
$$r(\tau) = \frac{\sum_{n=0}^{N-1} [x(n) - \mu_x][y(n+\tau) - \mu_y]}{\sqrt{\sum_{n=0}^{N-1} [x(n) - \mu_x]^2 \sum_{n=0}^{N-1} [y(n+\tau) - \mu_y]^2}} \quad (3.3)$$

kde x a y odpovídají jednotlivým signálům, τ značí posun a μ odpovídá střední hodnotě jednotlivých signálů. Níže popsany shlukovací algoritmus využívá Pearsonův korelační koeficient pro popis podobnosti mezi jednotlivými cykly. [14]

4 Rozhodování s využitím fuzzy logiky

V navrženém algoritmu je část, kde je potřeba aplikovat rozhodovací systém. Ten by měl na základě vybraných příznaků rozhodnout, zda shluk u kterého předpokládáme přítomnost pouze PVC tuto podmínku splňuje. Úkolem této části algoritmu je tedy odstranit z tohoto shluky cykly, nespádající do skupiny PVC. Místo rozhodování na základě pevně stanovených prahů byl zvolen postup využívající fuzzy logiku a fuzzy rozhodování. Tento postup mění pevné hranice prahů na neurčité (fuzzy), které - dle studií - v mnoha případech poskytují lepší výsledky.

Celý rozhodovací systém (fuzzy interferenční systém) založený na fuzzy logice můžeme rozdělit na několik dílčích částí. Popis jednotlivých cyklů spadajících do shluku je volen formou příznaků. Protože však v tomto případě pracujeme s neurčitými hranicemi je také potřeba pracovat s fuzzy prezentací příznaků. Z toho důvodu je nutné stanovit funkce příslušnosti pro **fuzzifikaci** (převod ostré číselné hodnoty příznaku na neurčitou hodnotu - tj. stupeň příslušnosti určité fuzzy množině). **Funkce příslušnosti** je funkce popisující stupeň příslušnosti jednotlivých prvků fuzzy množině (v tomto případě fuzzy množina představuje rozsah hodnot kterých mohou nabývat jednotlivé příznaky). Funkce může nabývat hodnot z intervalu $<0,1>$. Čím vyšších čísel funkce nabývá, tím více prvek spadá do konkrétní fuzzy množiny. Tvar funkce příslušnosti může být různý (viz Obrázek 4.1). [15] [16]



Obr. 4.1: Ukázka různých funkcí příslušnosti

Po fuzzifikaci jednotlivých příznaků následuje samotný proces rozhodování, ke kterému potřebujeme soubor pravidel (báze pravidel), na základě kterých bude systém rozhodování provádět. **Báze pravidel** je tvořena na základě zjištěných znalostí

o dané problematice. Pravidla jsou vyjádřena pomocí lingvistických proměnných a nabývají obecného tvaru:

$$\mathbf{POKUD} \text{ (} \textit{antecedent pravidla} \text{)} \rightarrow \mathbf{PAK} \text{ (} \textit{konekvent pravidla} \text{)},$$

kterému se říká zobecněné odvozovací pravidlo. **Antecedent** je konkrétní stav, který pozorujeme (v našem případě fuzzy hodnota příznaku daného cyklu) a **konekvent** je závěr pravidla, který vychází z konkrétního stavu. Důležité je, aby se každá vstupní proměnná (příznak) objevila alespoň v jednom antecedentu pravidla a aby pro každou výstupní proměnnou existovalo alespoň jedno pravidlo. Ideální však je, pokud jsou v bázi pravidel zahrnuty všechny možné stavy, které mohou nastat a navzájem se nevylučují. Pokud tedy mluvíme o fuzzy expertním systému, mluvíme o souboru několika Pokud-Pak pravidel.

Ve chvíli kdy je soubor pravidel vyjádřen pomocí lingvistických proměnných (slovně) je také potřeba tyto pravidla matematicky vyjádřit, aby bylo možné získat celkový výsledek. Závěr souboru Pokud-Pak pravidel je získáván pomocí **kompozičního pravidla úsudku**, kde jsou pravidla rozkládána a po částech řešen podle daných pravidel.

Pro implementaci těchto rozhodovacích pravidel existuje matematický aparát, ve kterém používáme různé operace (např. implikace, průnik, sjednocení) pro práci s fuzzy čísly a množinami fuzzy čísel. Podle způsobu provedení jednotlivých operací s fuzzy množinami rozeznáváme více druhů rozhodovacích systému, v tomto případě byl zvolen Mamdaniho rozhodovací systém.

Protože máme velké množství příznaků i samotných rozhodovacích pravidel je ve finále potřebné provést agregaci výstupů z dílčích Pokud-Pak pravidel do jedné výstupní hodnoty. Výstupní hodnotu rozhodovacího systému získáváme ve formě fuzzy množiny a pro získání konkrétního čísla je použita operace defuzzifikace (přiřazení fuzzy množině ostré hranice, které ji nejlépe reprezentují). [15] [17]

5 Popis analyzovaných dat

V této práci byla použita databáze patientských záznamů EKG z UBMI. Jednalo se o 6884 desetisekundových EKG záznamů snímaných z osmi svodů (II, III, V1-V6). Data byla snímána se vzorkovací frekvencí 500 Hz. U těchto dat již byla vyfiltrována stejnosměrná složka a 50 Hz brum. Spolu s databází byla možnost pracovat se souborem detekovaných pozic QRS komplexů a souborem anotací, vzniklém při ručním procházení záznamů na UBMI. Rozdělení v tomto souboru bylo provedeno do deseti neexkluzivních kategorií. Z těchto deseti kategorií osm odpovídalo jednotlivým patologiím, které mohou být v záznamu nalezeny, a zbylé dvě kategorie odpovídaly záznamům obsahující šum a záznamům naměřených u pacientů vlastní kardiostimulátor.

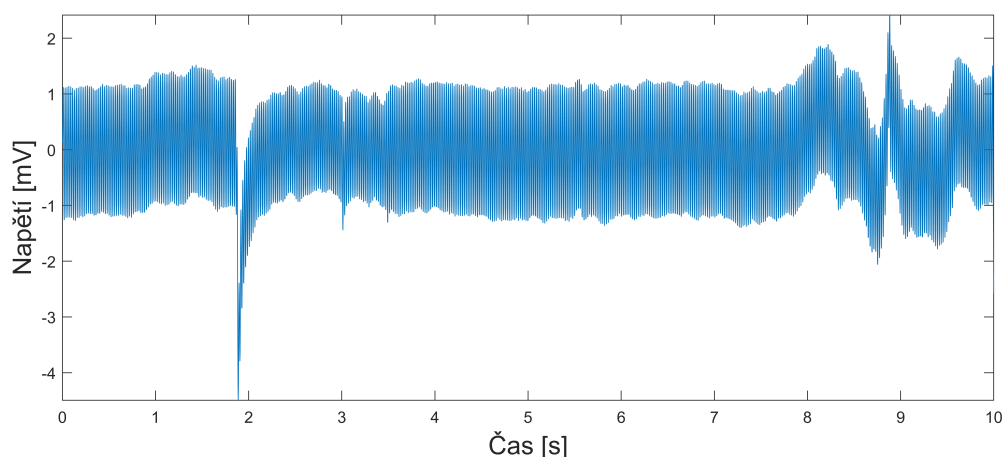
Z důvodu výskytu vysokonapětových impulsů kardiostimulátoru, zkreslujících následné zpracování, nebyla tato kategorie nadále zpracovávána. Databáze obsahovala také záznamy spadající pouze do kategorie se šumem, nebo záznamy nespádající do žádné z kategorií. Všechny výše zmíněné záznamy byly z databáze odstraněny. Po úpravě databáze obsahovala 6714 EKG záznamů. Pro představu je zde uvedena Tabulka 5.1 obsahující detailní přehled rozdělení do kategorií. V práci se však nepracovalo s celou databází. Pro zaručení objektivního hodnocení byla data rozdělena na trénovací a testovací tak, aby zastoupení patologií bylo v jednotlivých skupinách (trénovací a testovací) rovnoměrné. V průběhu návrhu a optimalizace byly používány trénovací data a testovací data byla použita pouze pro finální hodnocení.

Tab. 5.1: Přehled signálů zařazených do jednotlivých kategorií

Kategorie	Počet signálů celkem	Trénovací data	Testovací data
Atriální fibrilace	553	434	108
Sinusový rytmus	5942	4711	1178
Supraventrikulární rytmus	194	145	37
Ventrikulární rytmus	102	82	19
Ostatní	88	-	-
Atrioventrikulární blokáda	250	194	50
Blokáda Tawarových ramének	434	350	79
Síňová extrasystola	418	348	64
Komorová extrasystola	413	321	81
Zašuměné signály	2183	1664	415

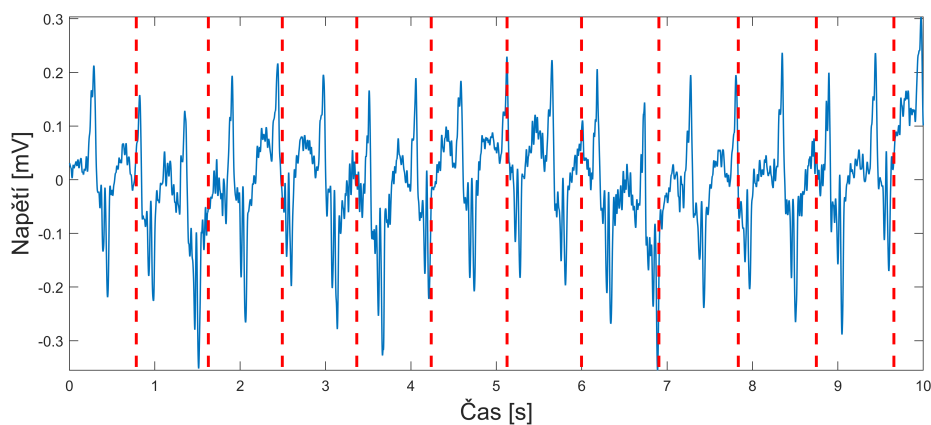
I přes odstranění výše zmíněných signálů, databáze stále obsahovala značně zašuměné signály. Příklad můžete vidět na Obrázku 5.1. V případě nalezení během

práce s programem, byly i tyto záznamy vyřazený z následné analýzy.

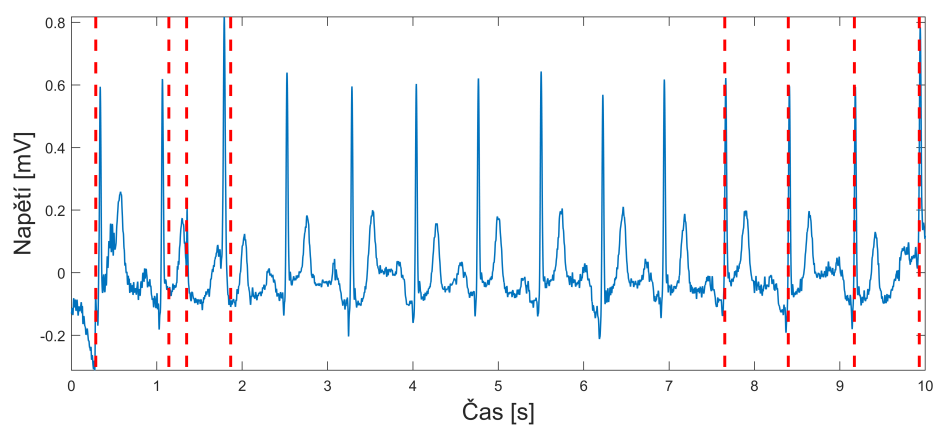


Obr. 5.1: Ukázka silně zašuměného signálu (svod II), dle anotace signál obsahující ABSV.

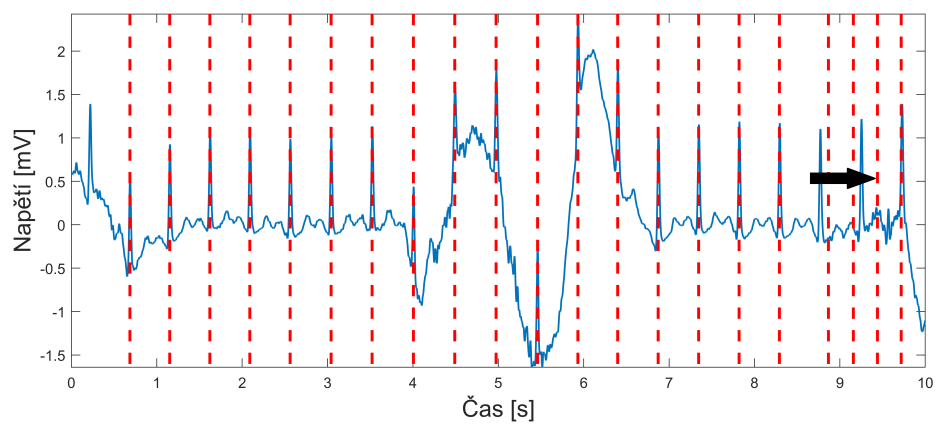
Ve výše zmíněném souboru detekovaných QRS komplexů byly nalezeny chyby v podobě nepřesné detekce vrcholů R kmitů (viz Obrázek 5.2), nedetekovaných komplexů (viz Obrázek 5.3), či detekce i v místech, kde se QRS komplex nenachází (viz Obrázek 5.4). V rámci detekce QRS komplexů je důležité zmínit, že nezdetekované QRS komplexy se často nachází právě na začátku a konci EKG záznamu a právě tyto chyby by mohly vést ke zhoršení výsledků následného třídění.



Obr. 5.2: Nepřesná detekce vrcholů R kmitů, signál č. 150, svod II.



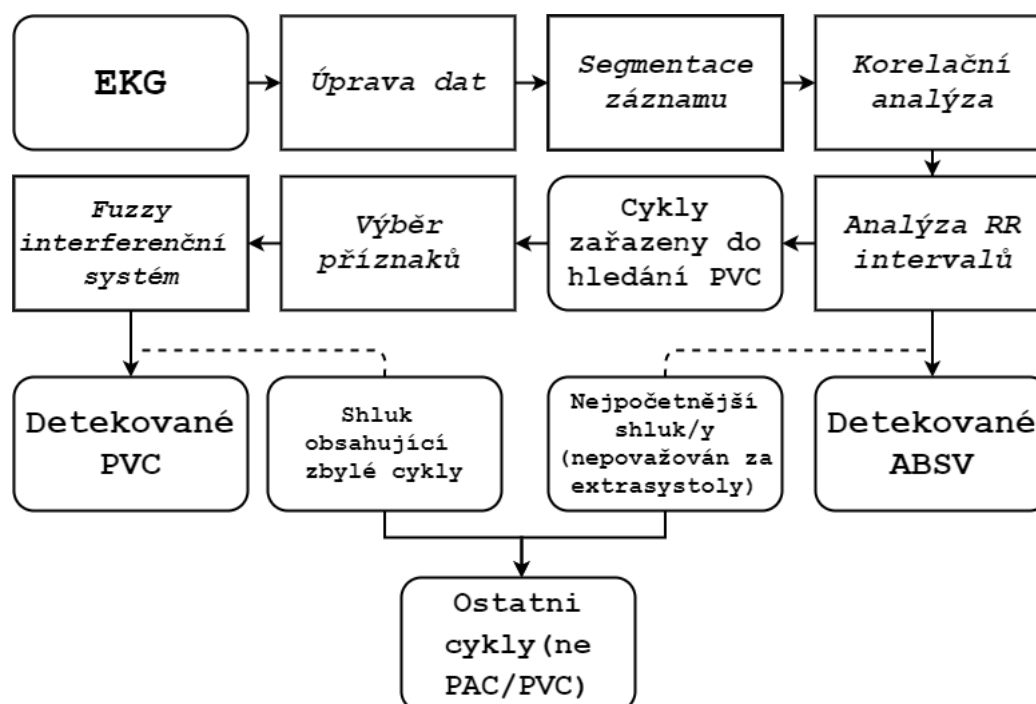
Obr. 5.3: Nedetekované QRS komplexy, signál č. 133, svod II.



Obr. 5.4: Detekce QRS komplexů i v místech, kde se komplex nenachází (místo označeno šipkou), signál č. 2051, svod II.

6 Programové řešení

Cílem této kapitoly je popis implementace vybraného algoritmu shlukování jednotlivých QRS komplexů ve vývojovém prostředí MATLAB R2019b. Algoritmus byl navržen pro detekci extrasystol v EKG záznamu a z toho důvodu je signál reprezentován souborem jednotlivých cyklů QRS a tyto cykly následně dělíme do tří skupin, kterými jsou komorové extrasystoly, sínové extrasystoly a zbylé cykly. Obecné schéma navrženého programového řešení je znázorněno na obrázku 6.1.



Obr. 6.1: Obecné schéma programového řešení.

6.1 Primární filtrace dat

Jak již bylo zmíněno v kapitole 5, součástí databáze byly i signály s chybně detekovanými pozicemi QRS komplexů a to i takové, že většina QRS komplexů detekována nebyla (viz Obrázek 5.3). Z toho důvodu se do algoritmu zahrnula funkce, která pouze na základě počtu QRS komplexů provádí jednoduché vyřazování takových cyklů. Pokud celkový počet QRS komplexů v tomto 10 s záznamu klesne pod hodnotu 5, což odpovídá tepové frekvenci 30 bpm, je z následného zpracování vyřazen.

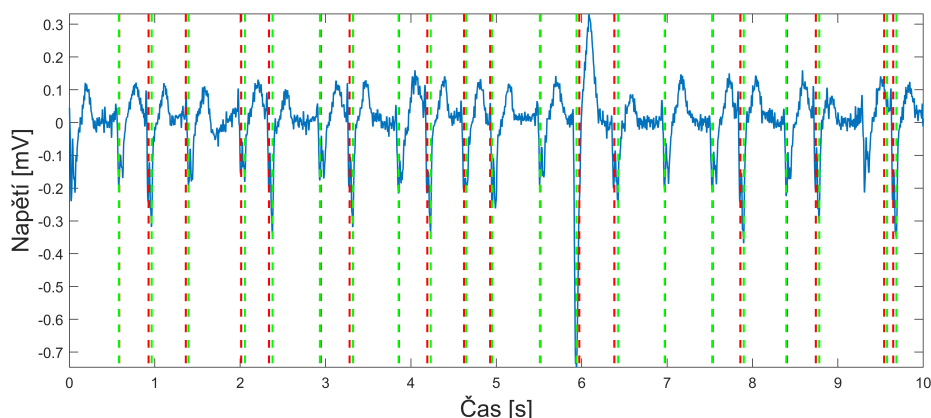
6.2 Optimalizace vrcholu R kmitu

Chyby v podobě nepřesných lokací vrcholů R kmitů, které se v souboru s pozicemi QRS komplexů také objevovaly řeší tato část programu. Protože kompletní snaha o úpravu všech chyb v souboru detekcí by byla velice náročná a úspěšnost by nebyla vysoká, bylo provedeno pouze upřesnění vrcholů, díky kterému by mělo dojít ke zkvalitnění následného zpracování, zvláště co se týče rozdělení signálu na cykly, korelace signálů a výběru příznaků.

Pro upřesnění pozic vrcholů bylo zvoleno okno stabilní velikosti v rámci celého desetisekundového záznamu. Jako vrchol R kmitu bylo považováno maximum v tomto okně. Tepová frekvence a tedy i počet QRS komplexů je v každém záznamu odlišný. Na tuto skutečnost musel být brán ohled při výběru počtu prvků okolí (velikost okna), ze kterých bude určováno maximum. Z toho důvodu bylo okolí zvoleno jako pětina průměrného počtu vzorků mezi jednotlivými vrcholy. Prvotní velikost okna byla třetina průměrného počtu vzorků, avšak v rámci optimalizace programu, kdy se formou iterativní optimalizace hledala velikost okna poskytující nejlepší výsledky byla tato hodnota změněna.

Aby program pracoval správně, musel být vyřešen problém R kmitů, které mají svoji maximální výchylku v záporných hodnotách. Z toho důvodu byla před samotným výběrem provedena primitivní analýza formou sumace hodnot vzorků na pozicích detekovaných R kmitů, která určovala většinové umístění vrcholů. V případě kladné sumy byl vrchol volen jako maximální kladná výchylka a v případě záporné sumy byl vrchol zvolen jako maximální záporná výchylka. Kromě této metody byla otestována také metoda založená pouze na počítání znamének (bez ukládání a sečítání hodnot maxim), která ovšem vykazovala horší výsledky.

Ve většině záznamů upřesnění vrcholů proběhlo správně (viz Obrázek 6.2). Je však nutné zmínit, že pokud jako vrchol byl označen útvar neodpovídající QRS komplexu, funkce přesto pouze vybrala maximum v okně. Dále je potřeba zmínit, že v některých záznamech nedosahovaly nejvyšších výchylek R kmity, ale jiné morfologické útvary. Ani v takovém případě nedošlo ke správné optimalizaci vrcholů.

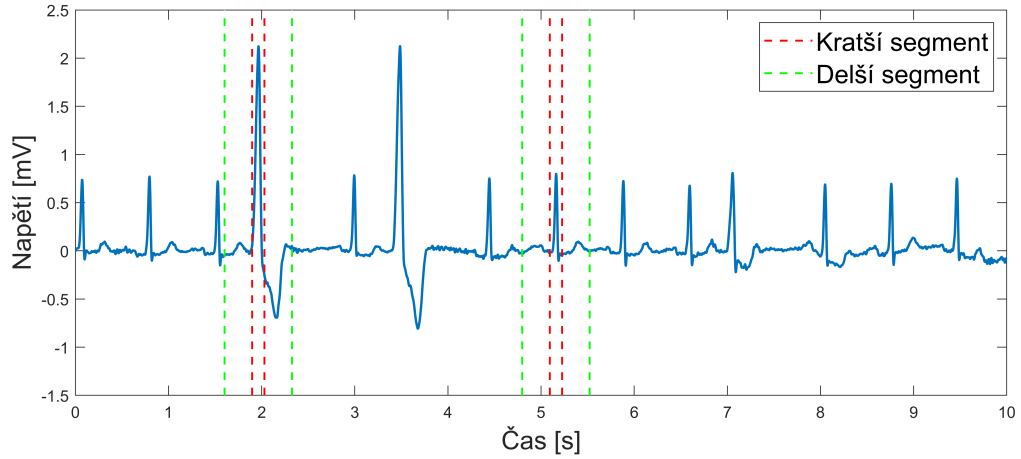


Obr. 6.2: Ukázka správné optimalizace pozic detekovaných QRS komplexů (signál č.5240, svod II), původní pozice (červeně), optimalizované pozice (zeleně). Na konci signálu lze vidět také optimalizaci pozice, která je chybně detekovaná. I zde se provedl pouze výběr minima (detekce vrcholu probíhá v záporných hodnotách).

6.3 Segmentace signálu

Při shlukové analýze se porovnává vzájemná podobnost jednotlivých cyklů a z toho důvodu bylo vhodné z EKG vybrat jednotlivé segmenty odpovídající těmto cyklům. I přes provedenou filtraci zůstaly některé signály zašuměné. Tento šum se z velké části nacházel mezi jednotlivými QRS komplexy. Pokud by se tyto části zahrnuly do korelační analýzy, výsledné shlukování by nebylo spolehlivé. Díky výběru pouze částí signálu kolem R kmitu bylo množství šumu zredukováno. Další problém, který musel být vyřešen byla volba délky jednotlivých segmentů. V případě výběru nedostatečně dlouhé oblasti, by nemusela být rozpoznána PVC, neboť by byla vybrána příliš krátká část, která s ostatními segmenty vysoce koreluje a odlišnost by byla přítomna mimo vybraný úsek (viz Obrázek 6.3). Proto byly z EKG vybrány dvě skupiny segmentů: skupina obsahující pouze R kmit a druhá skupina obsahující i jeho okolí a převážnou část cyklu. Pro další aplikace se pak vždy vybrala vhodnější skupina segmentů.

Na základě této myšlenky byla navržena funkce, která podle zadaných časových intervalů provede výběr segmentů. Časové intervaly jsou zadávány v ms a následně převedeny na počet vzorků od vrcholu kmitu R k vlně P a k vlně T. Volba časových intervalů vychází ze standardních hodnot intervalů EKG záznamu zdravého jedince (viz tabulka 1.1) a požadavku, aby skupina kratších segmentů obsahoval pouze QRS komplex. Pro maximální optimalizaci a přizpůsobení se jednotlivým záznamům byla při výběru časových intervalů zohledněna tepová frekvence (dále jen TF). Jako průměrný tep bylo vybráno 80 bpm, kde se bere segment dlouhý 100 ms



Obr. 6.3: Ukázka výběru kratšího a delšího segmentu u PVC (vlevo) a normálního cyklu (vpravo). Můžeme vidět nedostatečnou velikost kratšího okna v případě výskytu PVC.

(50 ms před R kitem a 50 ms za R kitem). Pro odlišné TF jsou délky kratších segmentů vypočítávány podle rovnice 6.1, kdy je segment zmenšováno pro nadprahovou TF, a naopak zvětšováno pro podprahovou TF. Níže zmíněná rovnice byla získána na základě analýzy výsledků shlukování s různými parametry intervalů.

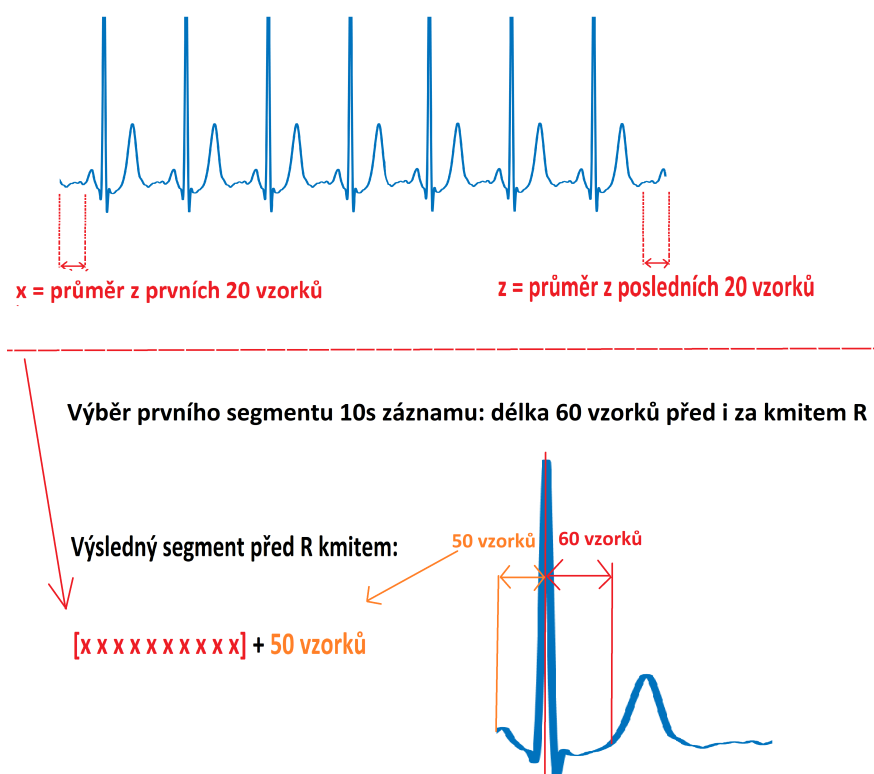
$$okno = 50 + |80 \pm TF| \cdot 0.1 \quad (6.1)$$

Velikost delších segmentů byla nastavena jako $\frac{1}{2}$ průměrného RR intervalu před i za R kitem. Ve výsledku je tedy délka delšího segmentu srovnatelná s délkou průměrného RR intervalu.

Funkce provede výběr segmentů na základě zadaných intervalů, které uloží do číselné matice a pošle na výstup. Ve výsledku získáme dvě skupiny segmentů - kratší a delší, které následně třídíme pomocí navrženého shlukovacího postupu.

Při výběru jednotlivých segmentů je použito okno délky $2N+1$, kde N je počet vzorků před a za detekovaným R kitem. Pro zdárnou segmentaci je tedy nutné splnit podmínku, že v okolí detekovaného R kmitu (jak před ním, tak za ním) je v signále dalších alespoň N vzorků. V případě nesplnění této podmínky - na začátku či na konci EKG - je zapotřebí provést doplnění daného úseku o potřebný počet vzorků. Podle pozice chybějících vzorků (konec či začátek EKG) se také určuje jejich hodnota. Pokud chybí n vzorků na konci, pak se jejich hodnota nastaví na hodnotu získanou zprůměrováním posledních 20 vzorků signálu (podobně i v případě začátku záznamu). Pokud je tedy QRS komplex detekován na padesátém vzorku a výsledný segment má mít délku 120 vzorků (60 vzorků před a 60 za R kitem), pak je

potřeba doplnit deset vzorků na začátek segmentu. Těchto deset vzorků má hodnotu průměrné hodnoty prvních dvaceti vzorků (viz Obrázek 6.4).



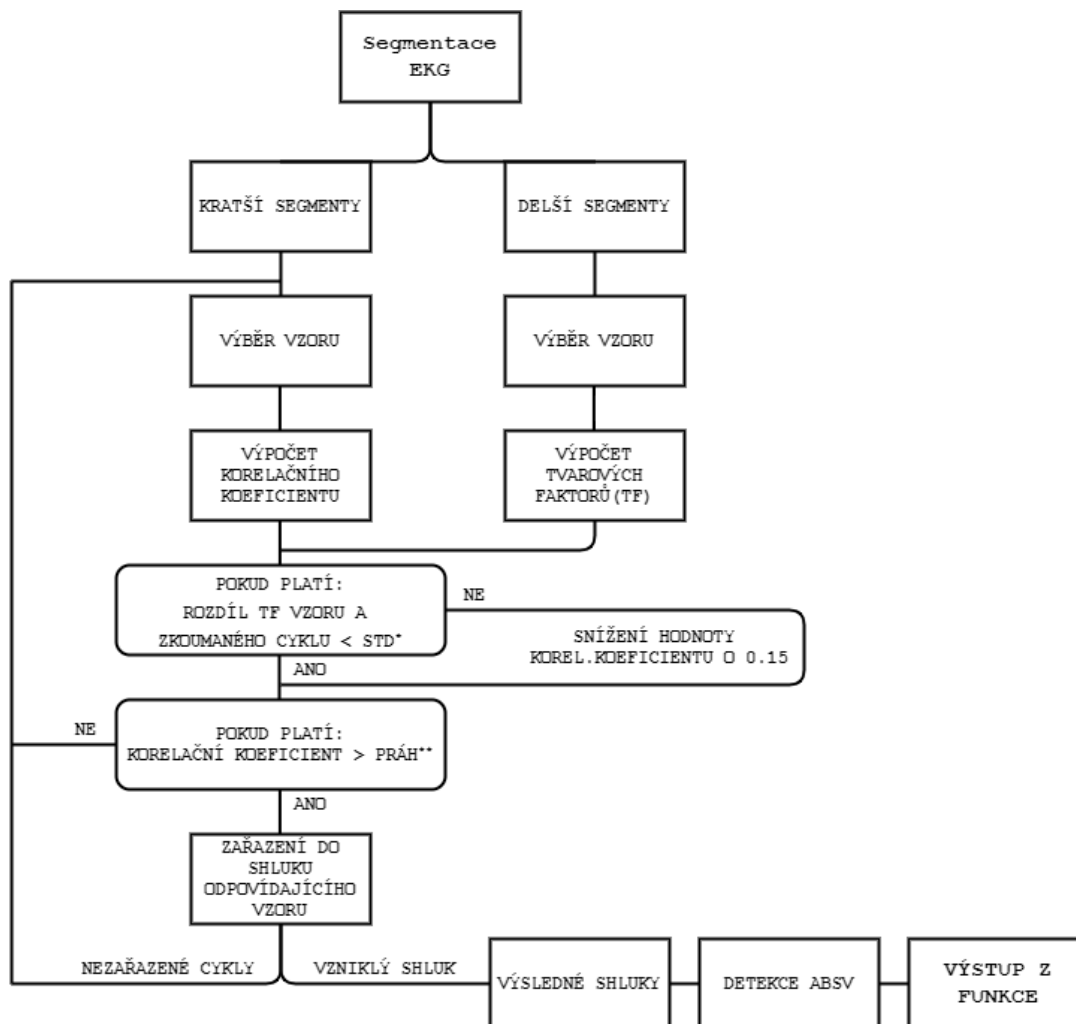
Obr. 6.4: Ukázka výběru segmentu při nedostatečném počtu vzorků na začátku signálu.

6.4 Shlukování na základě korelace

Všechny shlukové analýzy provádí samotné shlukování na základě vybrané vzdálenostní, či podobnostní metriky. Pro popis podobnosti jednotlivých cyklů za účelem shlukování byla v této práci vybrána korelační analýza. Důvod výběru právě tohoto přístupu byl zmíněn již v kapitole 3 a proto přistoupím rovnou k popisu navrženého algoritmu shlukování. Přehledové blokové schéma níže popsané funkce můžete vidět na Obrázku 6.5.

Na vstup této funkce přivádíme vektor RR intervalů daného záznamu a dvě skupiny segmentů. V rámci algoritmu je každá ze dvou skupin segmentů použita pro odlišné účely. Kratší segmenty, které by měli odpovídat pouze QRS komplexu byly využity pro korelační analýzu a delší segmenty byly použity pro získání tvarového

faktoru (viz kapitola 6.6). V tomto případě byly delší segmenty použity pro jejich vyšší vypovídající schopnost o odlišnosti cyklů. Při použití pouze krátkých segmentů docházelo k chybnému shlukování PVC z důvodu nerozpoznání odlišnosti vyskytující se mimo vybraný interval.



* STD - Hodnota směrodatné odchylky tvarových faktorů v rámci záznamu

** PRÁH - Hodnota korelačního koeficientu, kdy komplexy ještě zařazujeme do stejného shluku

Obr. 6.5: Přehledové schéma algoritmu shlukování na základě korelace a tvarového faktoru.

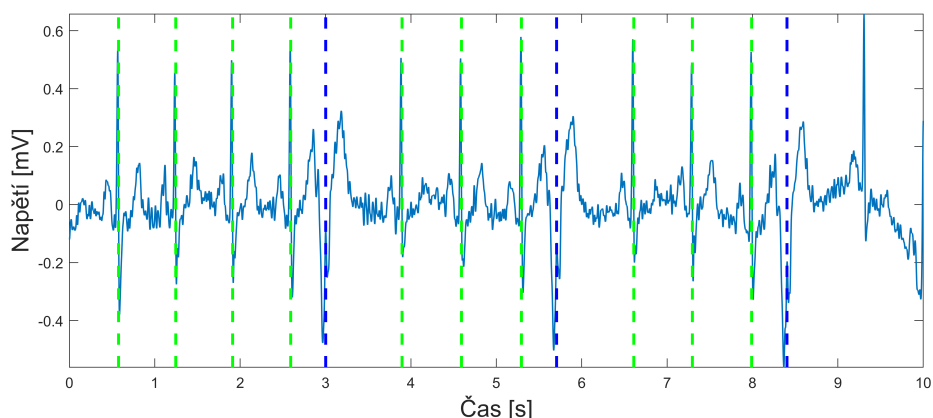
Korelační analýza je hlavní část programu, která provede prvotní třídění QRS komplexů v záznamu na základě velikosti korelačního koeficientu. Na začátku každé iterace se zvolí vzor – první QRS komplex (první kratší segment) – se kterým budeme zbylé komplexy (segmenty) korelovat. Funkce se vzorem prochází ostatní komplexy a u každého vypočítán Pearsonův korelační koeficient (viz kapitola 3).

Jakmile je vzor korelován se všemi zbylými komplexy, provede se samotné shlukování tak, že vzor představuje geometrický střed shluku a všechny komplexy, které mají korelační koeficient k tomuto vzoru vyšší než 0,84 jsou do tohoto shluku zařazeny. Hodnota 0,84 byla zvolena na základě optimalizačních postupů, kdy byl program opakovaně spouštěn, avšak vždy s jinou hodnotou a při prahové hodnotě 0.84 bylo na trénovacích datech dosaženo nejlepších výsledků. Nyní se vzor posouvá na další v pořadí nezařazený komplex. Iterace probíhají do doby, než jsou zařazeny všechny komplexy.

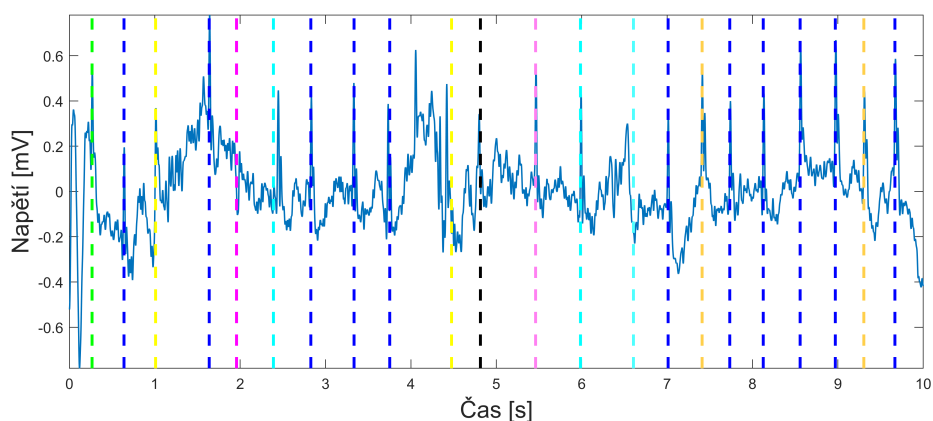
Pro přehlednost byl prvně popsán algoritmus shlukování podle korelačního koeficientu, avšak na rozhodování má vliv i hodnota tvarového faktoru vypočtená z delších segmentů. Ve chvíli provádění výpočtu korelačního koeficientu se hodnotí i rozdílnost tvarového faktoru mezi vzorem a analyzovaným komplexem. Pokud rozdíl tvarových faktorů přesahuje hodnotu směrodatné odchylky tvarových faktorů celého záznamu, automaticky se sníží hodnota korelačního koeficientu o hodnotu 0.15 a komplex je znevýhodněn při prahování. Hodnota 0.15 byla určena stejným způsobem jako práh v rámci korelace.

V této části programu byla také provedena volba svodu. Rozhodovací algoritmus je poměrně jednoduchý, jeho základní myšlenka je taková, že méně zašuměný signál bude rozřazen do menšího počtu shluků než signál zašuměný. Proto je celé shlukování provedeno zvlášť pro svody II, V4 a V6, které jsou pro analýzu EKG nejčastěji volené. Finální svod se volí takový, který má ve výsledku nejmenší počet vzniklých shluků. Předpoklad je takový, že vyšší procento šumu v analyzovaném EKG způsobuje snížení korelace, a tedy vznik více shluků, což je nežádoucí pro následnou analýzu cyklů.

Pokud je na vstup přiveden minimálně zašuměný signál, funkce přináší na výstup vhodně rozřazené cykly (viz Obrázek 6.6). Problém nastává u výrazně zašuměných signálů, v takovém případě se cykly rozřadí do většího počtu shluků, než je potřeba (viz Obrázek 6.7).



Obr. 6.6: Ukázka výsledku shlukování na základě korelace v případě minimálně zašuměných dat, různou barvou jsou od sebe odděleny vzniklé shluky. Z vizuální analýzy můžeme usuzovat, že shluk označený modře reprezentuje PVC cykly a shluk označený zeleně cykly Ostatní.



Obr. 6.7: Ukázka výsledku shlukování na základě korelace v případě vysoce zašuměných dat, dle barvy patrné rozřazení do mnoha shluků.

6.5 Detekce síňových extrasystol

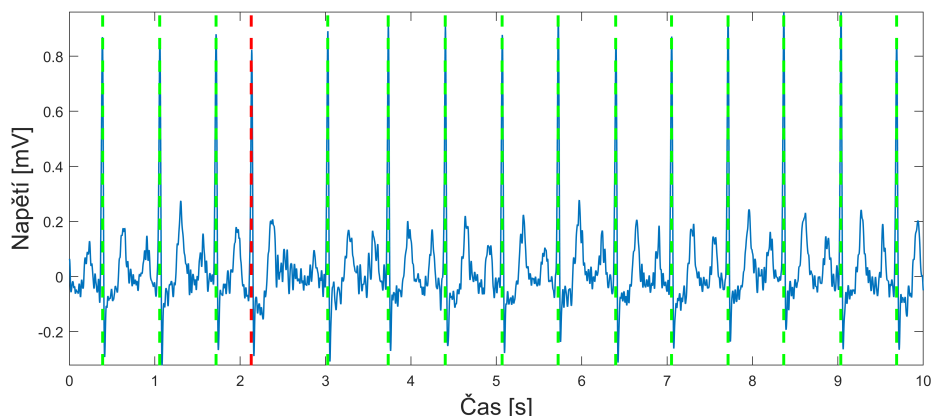
Výše popsáný postup však neřeší detekci PAC. K tomu slouží další implementovaný postup, založený na rozpoznání odlišnosti RR intervalů charakteristických pro daný typ extrasystol. Vstupem je vektor RR intervalů konkrétního záznamu, konstanta označující nejpočetnější shluk/y (právě v tomto/těchto probíhá detekce PAC) a vektor vypovídající o výsledku prvotního shlukování (vektor čísel, přičemž každé číslo odpovídá jednotlivému shluku). Detekce probíhá v nejpočetnějších shlucích právě z toho důvodu, že v případě PAC by se QRS komplex neměl lišit od ostatních

(kromě PVC) a korelace by tyto cykly měla zařadit do stejného shluku. Selhání algoritmu pro případ extrémně zašuměných dat, kdy se PAC nemusí dostat do nejpočetnější skupiny, nebo se do této skupiny dostane PVC, se v rámci dané bakalářské práce nepodařilo ošetřit. Na výstupu poté očekáváme upravený vektor vypovídající o rozřazení do shluků (přidání shluku pro PAC), přičemž skupina komplexů odpovídající shluku pro PAC má vždy označení 1.

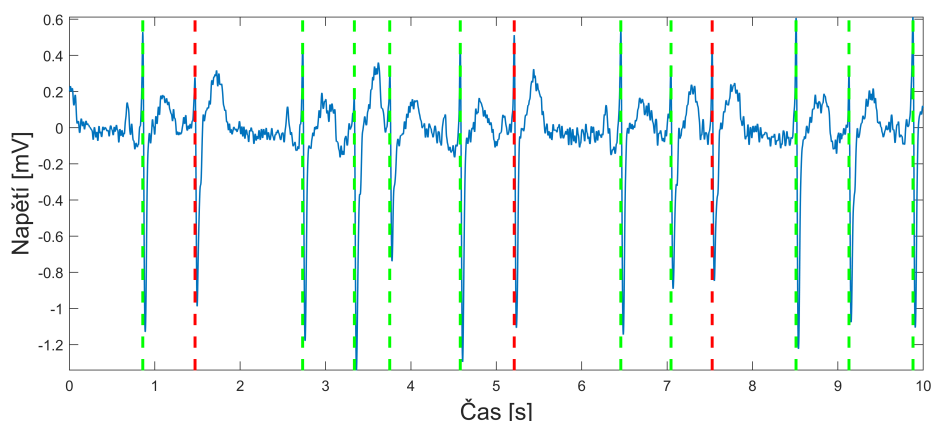
Algoritmus detekce PAC je založen na skutečnosti, že u PAC je RR interval před extrasystolou zkrácený a naopak RR interval za proběhlou extrasystolou je prodloužený z důvodu výskytu neúplné kompenzační pauzy. Algoritmus ošetřuje i případy výskytu následných mnohočetných extrasystol, kde předpokládáme kaskádu krátkých RR intervalů ukončených neúplnou kompenzační pauzou. Pokud do algoritmu vstoupí cyklus, který má kratší RR interval (již brán za extrasystolický) před i za QRS komplexem, funkce iterativně pokračuje v kontrole RR intervalů následných cyklů. Sada těchto cyklů se vyhodnotí jako extrasystolická pouze v případě, pokud po posledním cyklu v řadě krátkých RR intervalů před i za QRS komplexem následuje cyklus s krátkým RR intervalem před komplexem a delší pauzou za komplexem QRS (interval brán již za postextrasystolický).

Obdobnými metodami, kterými byl zvolen úbytek korelačního koeficientu nebo korelační práh, byly zvoleny prahy, kdy je RR interval brán již za extrasystolický. Výchozími orientačními body během optimalizace byly střední hodnoty extrasystolických a postextrasystolických RR intervalů v rámci trénovacích dat. Opět je zde zohledněna TF, a proto byl zvolen práh zvlášť pro bradykardii, tachykardii a fyziologickou TF. Je monitorován jak extrasystolický RR interval (bere se jako extrasystolický, pokud je jeho hodnota nižší než hodnota průměrného RR intervalu v záznamu zmenšená o hodnotu prahu) tak i postextrasystolický interval (k průměrnému RR intervalu práh přičítáme a RR interval musí být větší, aby byl považován za postextrasystolický). Pokud je tedy signál vysoce podobný s ostatními, ale liší se od průměrného RR intervalu v rámci záznamu o více jak 50/80/130 ms v závislosti na TF (RR interval před komplexem je menší než průměrná hodnota - 50/80/130 ms a RR interval za komplexem je delší než průměrná hodnota + 50/80/130 ms) vyhodnotí se jako PAC. Ukázkou správné funkce detekce PAC můžeme vidět na Obrázku 6.8, kdy se výsledky shlukování shodují s expertním hodnocením daného EKG záznamu.

I přes snahu o maximální optimalizaci funkce existují arytmie, u kterých dochází k vysokým počtům falešně pozitivních výsledků. Jedná se převážně o supra-ventrikulární arytmie, kde dochází k atrio-ventrikulární disociaci, což zapříčiňuje nepravidelné RR intervaly, které jsou při analýze problematické (viz Obrázek 6.9).



Obr. 6.8: Ukázka správné detekce PAC, dle vizuální analýzy zeleně značený shluk Ostatní a červeně nově zdetekované PAC



Obr. 6.9: Ukázka selhání algoritmu v případě atrio-ventrikulární disociace rytmů (chybně detekované PAC označené červeně).

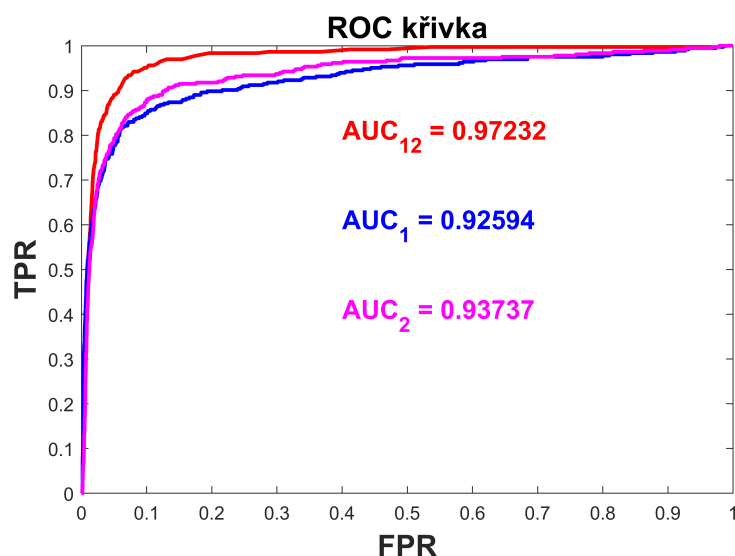
6.6 Výběr příznaků

Nyní již máme samostatný shluk pro PAC a nejpočetnější shluk/y, u kterých nepředpokládáme výskyt PVC kvůli jejich výrazně odlišné morfologii. Do další části programu vstupují zbylé shluky, v nichž proběhne detekce PVC na základě vybraných specifických EKG příznaků a vyhodnocení pomocí navrženého fuzzy expertního systému.

Do funkce pro výběr příznaků vstupují všechny cykly (dlouhé i krátké), avšak spolu s nimi přivádíme na vstup číselný vektor označující pozice cyklů zahrnutých do další analýzy. Pouze u označených cyklů probíhá extrakce příznaků. Dále vstupuje vektor RR intervalů a vektor představující referenční cyklus pro daný desetisekun-

dový záznam. Tento cyklus byl získán zprůměrováním cyklů, které již nevstupují do této části algoritmu a tudíž předpokládáme že se jedná o fyziologické cykly.

V rámci optimalizace této části algoritmu se hledala kombinace příznaků, která by dosahovala nejlepších výsledky. Celkem bylo vyzkoušeno kolem 40 různých EKG příznaků z časové oblasti a z nich bylo vybráno 6 nejvhodnější. Selektce příznaků probíhala jak formou pokus-omyl tak i pomocí vizualizace boxplotů a hodnocení selektivních vlastností jednotlivých příznaků. V rámci hodnocení byly vždy vykresleny dva boxploty vedle sebe pro dvě skupiny cyklů (PVC a ostatní). Pokud mezi zobrazenými boxploty existoval práh dělící tyto data jednalo se o vhodný příznak. Ve chvíli, kdy byla vybrána určitá podskupina příznaků byla provedena ROC analýza pro kvalitnější určení jejich separačních vlastností. Finálním krokem těchto postupů bylo vybrání 6 příznaků, které jsou podrobně popsány níže. Původně bylo vybráno 7 příznaků, avšak později bylo zjištěno, že spojením dvou příznaků popisujících napěťové rozdíly signálů, získáme jeden silnější příznak (viz Obrázek 6.10). Tento příznak byl získán podílem dvou původních příznaků (viz Obrázek 6.11). Boxploty a výsledky ROC analýzy vybraných příznaků můžeme vidět v Příloze A. V rámci ROC analýzy můžeme vidět hodnotu AUC (angl. area under curve), která nám také určuje kvalitu příznaků. Čím více se tato hodnota blíží číslu 1, tím je příznak kvalitnější.



Obr. 6.10: Ukázka zvýšení hodnot v rámci ROC analýzy při spojení dvou příznaků v jeden. AUC_1 - AUC pro první příznak popisující rozdíl maximálních výchylek segmentu a reference, AUC_2 - AUC pro druhý příznak popisující rozdíl minimálních výchylek segmentu a reference, AUC_{12} - AUC výsledného příznaku získaného poměrem dvou předchozích.

Výstupem této funkce je matice příznaků, která slouží pro detekci PVC prováděnou v dalších krocích. V práci se pracovalo s níže popsány příznaky.

RR interval

Vzdálenost mezi sousedními R kmity v signálu (RR interval) patří mezi základní příznaky při detekci srdečních arytmií a tedy i komorových extrasystol. Hlavním důvodem, proč RR interval dobře popisuje srdeční arytmiu je, že odchylky vnímané v šířce intervalů RR jsou korelovány s odchylkami v morfologii křivky, často vyvolané přítomností arytmií.

V práci byl jako jeden z příznaků vybrán podíl RR intervalů před a za QRS komplexem a to z toho důvodu, že v rámci PVC nedochází pouze ke zkracování RR intervalu před extrasystolou, ale i k prodloužení RR intervalu za ní (viz kapitola 1). Zatímco u fyziologických cyklů by se měl tento podíl pohybovat kolem hodnoty 1, u extrasystol bude nabývat hodnot nižších.

Protože změna délek RR intervalů je pro extrasystoly výrazným příznakem, byl vybrán ještě jeden příznak popisující jejich odlišnost. Tímto příznakem je rozdíl RR intervalu před komplexem QRS a průměrnou hodnotou RR intervalů v daném záznamu EKG.

Rozdíl tvarových faktorů

Pro zdůvodnění výběru druhého příznaku je potřeba nejprve zavést pojem tvarový faktor (form factor, FF). **Tvarový faktor** je jeden ze tří Hjorthových deskriptorů, které jsou zvláštní tím, že charakterizují spektrální vlastnosti signálu na základě dat získaných v časové oblasti. Prvním z nich je aktivita $A(x)$, která zhruba odpovídá střednímu výkonu signálu a jde vlastně o hodnotu rozptylu σ_x^2 pro daný úsek signálu. Dále je mobilita $M(x)$ (6.2), která vypovídá o střední frekvenci signálu a vypočítá se jako odmocnina podílu aktivity první derivace signálu ku originální aktivitě.

$$M(x) = \sqrt{\frac{\sigma_{x'}^2}{\sigma_x^2}} \quad (6.2)$$

Posledním je již zmíněný tvarový faktor, nebo také složitost FF (6.3), která je odhadem šířky pásma signálu. FF je vypočítán jako odmocnina rozdílu mobility originálního signálu a mobility první derivace signálu. [14]

$$FF = M_{x'} - M_x = \sqrt{\frac{\sigma_{x''}^2}{\sigma_{x'}^2} - \frac{\sigma_{x'}^2}{\sigma_x^2}} \quad (6.3)$$

V práci nebyl použit samotný tvarový faktor, ale absolutní hodnota rozdílu mezi FF analyzovaného komplexu a referenčním komplexem, který reprezentuje fyziologický tvar QRS komplexu daného záznamu. V případě analýzy PVC by tato hodnota

měla být větší než 0, protože ve většině případů nabývají hodnoty FF u PVC vyšších hodnot. V ostatních případech by se měla pohybovat kolem hodnoty 0.

Obsah plochy pod křivkou

Dalším vybraným příznakem byla plocha pod křivkou jednotlivých QRS komplexů. Každý komplex má díky svým vlnám a kmitům určitou velikost plochy pod křivkou. Podle základních charakteristik víme, že PVC se vyznačují rozšířeným komplexem a tím pádem zvětšením této hodnoty. Žádoucí je, aby příznak obsahoval spíše informaci o rozdílnosti těchto ploch mezi PVC a zbytkem, a z toho důvodu byl jako příznak zvolen podíl ploch zkoumaného komplexu a již zmíněné reference delších segmentů, obsahující větší část cyklu.

Počet vzorků mezi extrémy

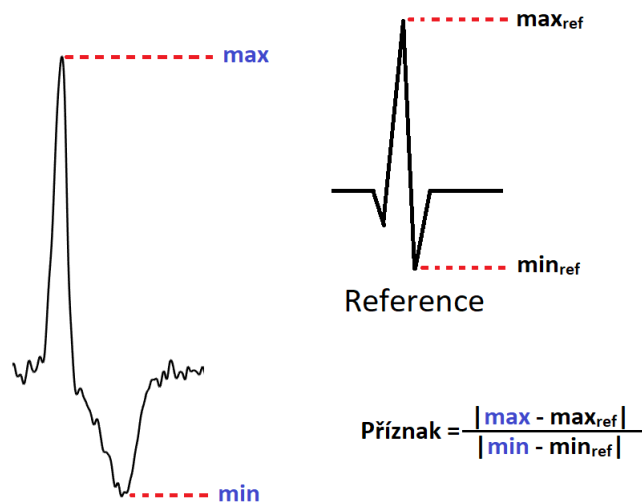
Protože práce neobsahuje detailní rozměření signálu, byla snaha nějakým způsobem kvantifikovat šířku jednotlivých komplexů. Toho bylo docíleno výpočtem vzorků mezi maximální a minimální hodnotou každého komplexu, čímž byla nahrazena informace o šířce komplexu.

Rozdíly napěťových špiček

Jeden z dalších často pozorovaných jevů v rámci PVC byla jejich napěťová odlišnost. Popis tohoto jevu byl proveden pomocí jednoho příznaku který byl vypočítán jako podíl rozdílu maximálního napětí kratšího segmentu a reference ku rozdílu minimálního napětí kratšího segmentu a reference (viz Obrázek 6.11), který byl zmíněn výše.

6.7 Detekce komorových extrasystol na základě fuzzy usuzování

Finálním krokem je detekce komorových extrasystol pomocí fuzzy expertního systému. Při návrhu tohoto programu byla trénovací data rozdělena do tří skupin, konkrétně se jedná o PAC, PVC a ostatní. V následujících krocích bylo pracováno pouze se skupinami PVC a ostatní. V rámci skupiny PVC byla potřeba doplnit anotaci o přesné pozice těchto extrasystol. Protože rozpoznat, zda se jedná o PVC či ne bylo mnohdy obtížné, nebyly anotovány všechny PVC. Vybrány byly pouze ty cykly u kterých bylo jisté, že se jedná o PVC. Celkem bylo nalezeno 364 komorových extrasystol u kterých proběhl výběr příznaků. Také u skupiny ostatní proběhl výběr příznaků, těchto cyklů bylo celkem 14602.

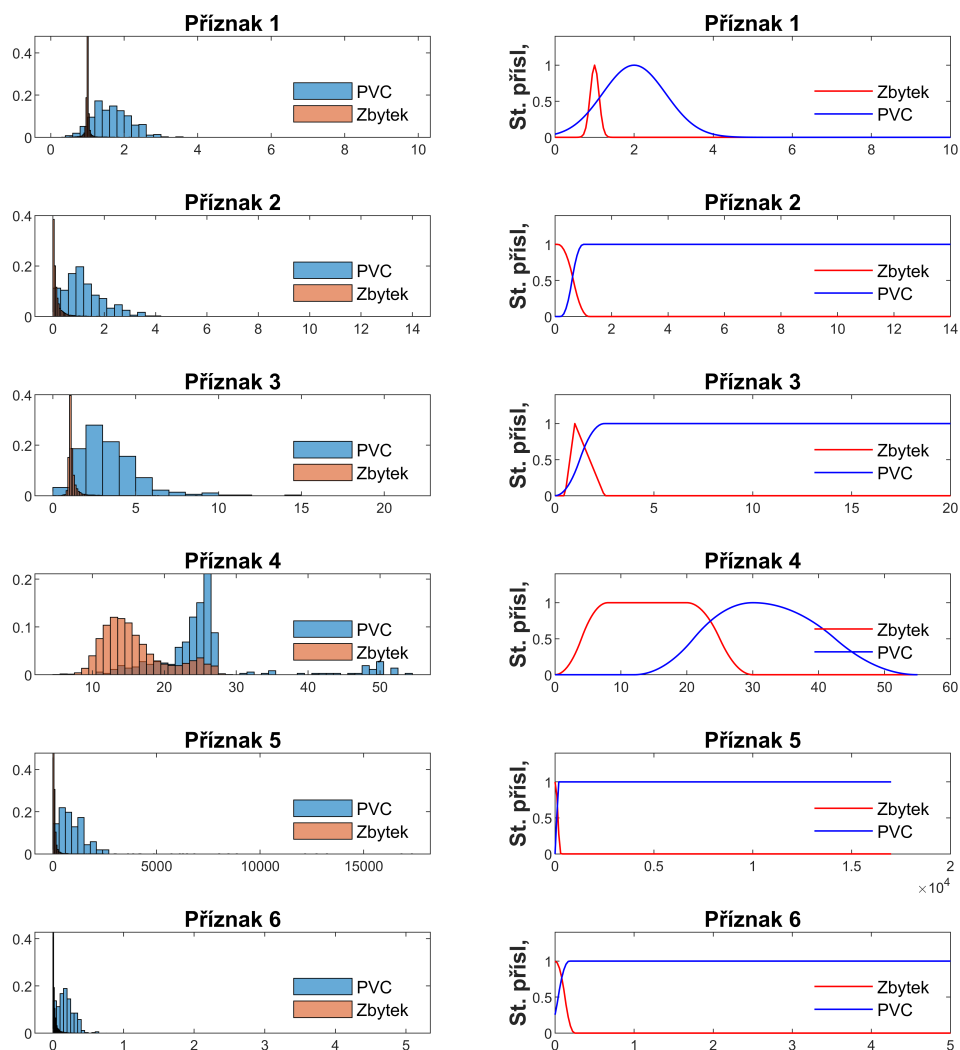


Obr. 6.11: Ukázka získání příznaku spojením dvou původních příznaků, čímž byl získán silnější příznak. Jedná se o rozdíl napětových špiček segmentu a reference podělený rozdílem napětových minim segmentu a reference.

První přístup volen v tomto kroku algoritmu byl založen na hledání prahu pro každý příznak pomocí ROC analýzy a následné vyhodnocování formou kontroly, zda příznak nabývá nadprahových hodnot. Pokud byly nadprahové hodnoty u více jak 5 příznaků, byl cyklus zařazen mezi PVC. V kapitole 4 jsme si mohly přečíst, že využitím fuzzy logiky se více přibližujeme lidskému rozhodování. Právě toto (navrhnout algoritmus, který by se více podobal postupu lékařů při hodnocení EKG) byl jeden z cílů této práce. Na základě těchto informací byl výše popsán postup nahrazen fuzzy expertním systémem typu Mamdani. Prvotní implementace fuzzy systému přispěla k získání lepších výsledků detekce PVC, než v případě jednoduchého prahování a proto byl v algoritmu ponechán a optimalizován.

Jeden z hlavních kroků při návrhu fuzzy expertního systému je návrh jednotlivých funkcí příslušnosti. Funkce příslušnosti byly zvoleny na základě analýzy histogramů příznakových hodnot (viz Obrázek 6.12). Systém obsahuje celkem sedm vstupních proměnných (každému příznaku odpovídá jedna vstupní proměnná) a jednu výstupní (viz Obrázek 6.13), přičemž pro každou proměnnou jsou navrženy dvě funkce příslušnosti (jedna pro PVC cyklus a druhá pro cykly neodpovídající PVC).

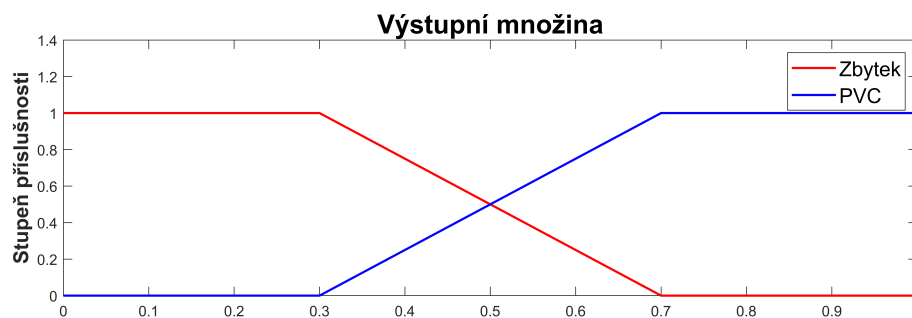
Dalším krokem při návrhu FIS bylo vytvoření báze pravidel. V bázi pravidel by měly být zahrnuty všechny stavy, které mohou nastat. Pokud tedy máme 6 příznaků, přičemž v rámci každého příznaku může nastat jeden ze dvou stavů (spadá do PVC či do skupiny ostatní) pak je počet pravidel, potřebných pro úplné pokrytí



Obr. 6.12: Návrh funkci příslušnosti pro jednotlivé vstupní proměnné rozhodovacího systému: vlevo - histogramy příznaků, vpravo - odpovídající funkce příslušnosti.

2⁶. Protože by tvorba tolika pravidel ručně zabrala spoustu času bylo využito funkce z Deep Learning Toolbox. Tato funkce vytváří všechny možné kombinace vektorů. Na vstup zadáme pouze čísla, která se mohou dosazovat na jednotlivé pozice a výstupem je matice vektorů obsahující všechny možné kombinace. V našem případě může být na každé pozici {1;2}.

Pravidla byla koncipována podobně jako v případě prahování a to tak, že pokud 5 a více příznaků odpovídá PVC cyklu pak je vyhodnocen jako komorová extrasystola. Pro ukázkou uvádím dvě pravidla, ale pro přehlednost pouze se třemi příznaky:

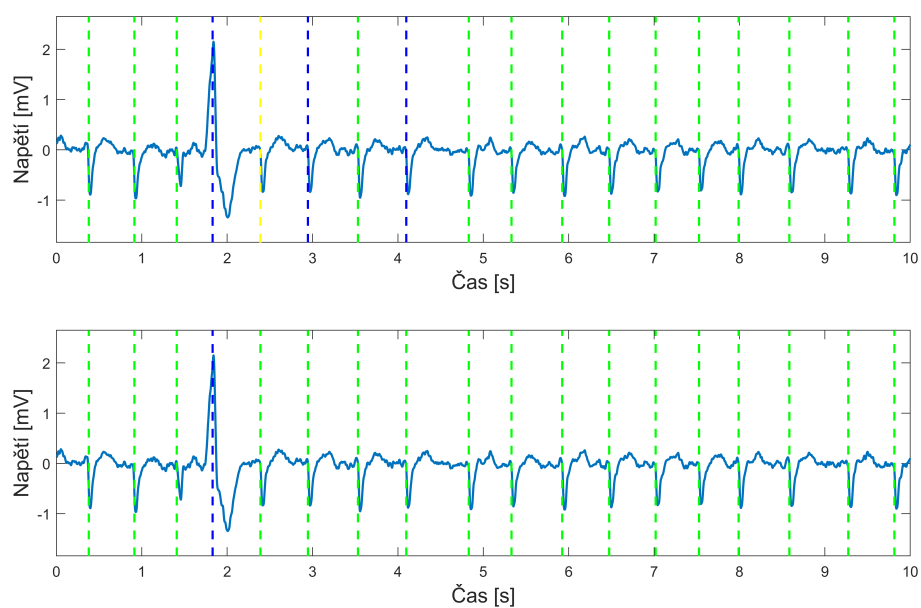


Obr. 6.13: Výstupní funkce příslušnosti. Červeně funkce příslušnosti pro ostatní cykly a modře pro PVC.

IF pr1 is PVC and pr2 is PVC and pr3 is PVC **THEN** výsledek is **PVC**

IF pr1 is PVC and pr2 is ostatní and pr3 is ostatní **THEN** výsledek is **OSTATNÍ**

Toto byl poslední krok nastavení systému. Navržený systém se nyní použije ve funkci pro vyhodnocení PVC. Na vstup funkce přivádíme matici příznaků, navržený systém a číselný vektor udávající informaci, které cykly budou vyhodnocovány. Samotné vyhodnocení probíhá pomocí FIS na jehož vstup přivádíme postupně příznaky pro jednotlivé cykly a na výstupu očekáváme hodnotu v rozsahu $<0,1>$. Cyklus je vyhodnocen jako PVC (přiřazení čísla 3 do vektoru odpovídajícímu rozřazení do shluků) pouze pokud je tato hodnota vyšší než 0.5, v opačném případě je zařazen do shluku Ostatní (přiřazení čísla 2). Výstupem této funkce, a také finálním výsledkem programového řešení je číselný vektor s možnými číslicemi 1 (viz kapitola 6.5), 2, 3. Každá číslice označuje jednotlivé shluky, zleva doprava to jsou PAC, Ostatní a PVC. Ukázkou správné funkce můžete vidět na Obrázku 6.14.



Obr. 6.14: Ukázka správné funkce programu pro detekci PVC. Prvotní shluky po shlukování (nahore), konečné rozdělení cyklů po úpravě pomocí fuzzy systému (dole), dle anotace modře označený shluk PVC a zeleně shluk Ostatní.

7 Vyhodnocení algoritmu

Finální částí praktické části je statistické vyhodnocení. Výsledky budou uvedeny pro trénovací i testovací data a to samostatně pro detekce PAC, PVC ale i výsledek pro detekci extrasystol obecně. Jak již bylo uvedeno v kapitole popisující analyzovaná data, skupina trénovacích dat obsahovala celkem 5354 záznamů a testovací data 1342 záznamů EKG. Ve výsledku tedy bylo analyzováno 6696 záznamů EKG. První dvě kontingenčních tabulky informují o kvalitě shlukování pomocí následujících parametrů:

TP (true positive, pravdivě pozitivní) – dle anotace záznam s přítomností PAC/PVC zařazen do shluku PAC/PVC.

TN (true negative, pravdivě negativní) – dle anotace záznam bez výskytu PAC/PVC zařazen do shluku Ostatní.

FP (false positive, falešně pozitivní) – dle anotace záznam bez výskytu PAC/PVC zařazen do shluku PAC/PVC.

FN (false negative, falešně negativní) – dle anotace záznam s přítomností PAC/PVC zařazen do shluku Ostatní.

Tab. 7.1: Tabulka detekce PAC a PVC pro trénovací data

	PAC	PVC
TP	291	237
TN	4490	4760
FP	516	274
FN	57	83

Tab. 7.2: Tabulka detekce PAC a PVC pro testovací data

	PAC	PVC
TP	47	62
TN	1159	1185
FP	116	73
FN	17	19

Z těchto hodnot byly následně odvozeny další statistické parametry, konkrétně **senzitivita** - úspěšnost nalezení extrasystoly při jejím výskytu (7.1), **specifita** - správnost nezařazení do shluku pro extrasystoly v případě záznamu bez extrasystoly (7.2), **pozitivní prediktivní hodnota** - pravděpodobnost s kterou cyklus

zařazený do skluku pro extrasystoly opravdu odpovídá extrasystole (7.3), **přesnost** - celková pravděpodobnost úspěšného shlukování (7.4) a **F1 skóre** (7.5), které můžeme považovat za harmonický průměr mezi PPV a TPR (vhodná volba v případě nerovnoměrného zastoupení jednotlivých kategorií).

$$TPR = \frac{TP}{TP + FN} \quad (7.1)$$

$$FPR = \frac{TN}{TN + FP} \quad (7.2)$$

$$PPV = \frac{TP}{TP + FP} \quad (7.3)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (7.4)$$

$$F = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} \quad (7.5)$$

Tab. 7.3: Výsledky detekce PAC

Data	TPR	FPR	PPV	ACC	F1
Trénovací	83.60 %	89.70 %	36.06 %	89.30 %	50.40 %
Testovací	73.40 %	91.00 %	29.00 %	90.00 %	41.40%

Tab. 7.4: Výsledky detekce PVC.

Data	TPR	FPR	PPV	ACC	F1
Trénovací	74.00 %	94.60 %	46.40 %	93.30 %	57.00 %
Testovací	76.50 %	94.20 %	45.90 %	93.10 %	57.40 %

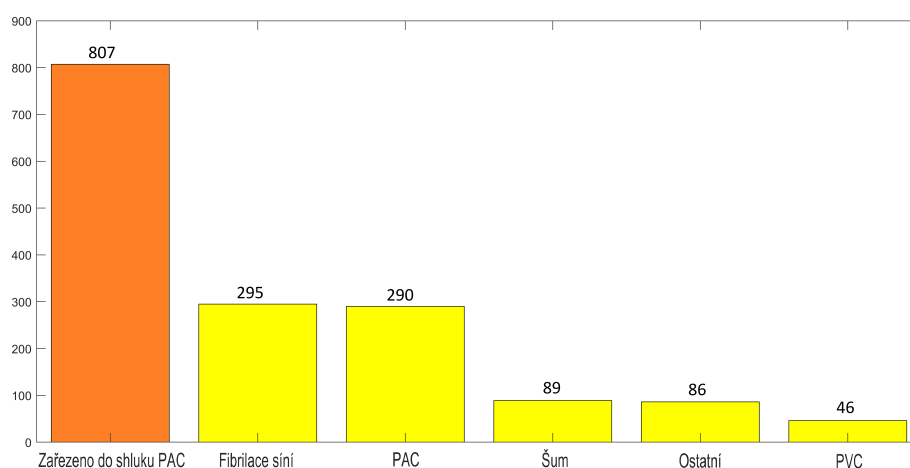
Tab. 7.5: Souhrnné výsledky detekce extrysystol získané zprůměrováním předchozích dvou tabulek.

Data	TPR	FPR	PPV	ACC	F1
Trénovací	78.80 %	92.10 %	46.40 %	91.30 %	53.70 %
Testovací	75.00 %	92.60 %	37.40 %	91.60 %	49.40 %

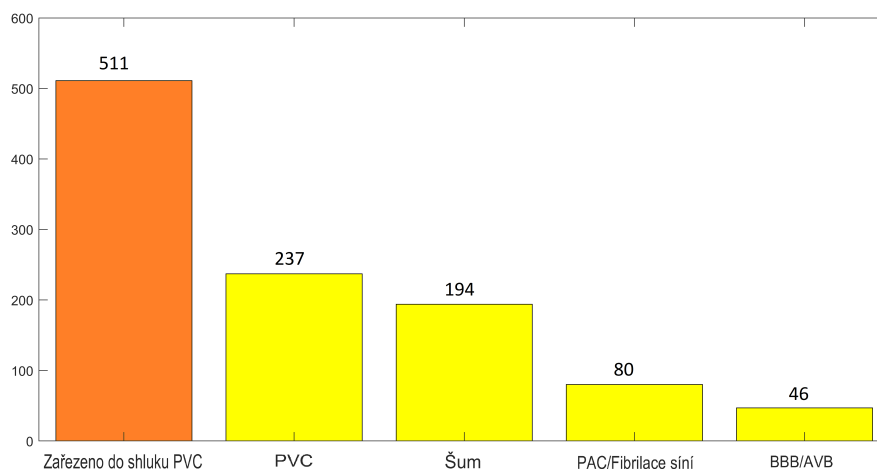
Z výsledků můžeme usuzovat, že v případě nerovnoměrného zastoupení jednotlivých kategorií jsou vhodnější parametry PPV a F1, které jsou citlivější na vysoké hodnoty falešně pozitivních detekcí. Pokud by byly pro hodnocení algoritmu použity pouze parametry TPR, FPR a ACC chybně by to poukazovalo na úspěšnější klasifikátor.

Celková úspěšnost navrženého klasifikátoru nedosahuje příliš kvalitních výsledků a to z důvodu vysokého počtu falešně pozitivních detekcí. Nízké hodnoty F1 a PPV jsou v případě PAC způsobeny hlavně přítomností fibrilace síní (viz Obrázek 7.1). V případě PVC se jedná převážně o přítomnost šumu v záznamu EKG (viz Obrázek 7.2).

V případě detekce PVC se výsledky na trénovací i testovací sadě dat shodují, avšak v případě PAC se některé parametry (TPR, F1) liší skoro o 10 %. Odlišné výsledky by mohly znamenat rozdílné zastoupení jednotlivých kategorií, nebo odlišnou kvalitu dat v rámci trénovací a testovací sady, což není přípustné.



Obr. 7.1: Sloupcový graf zobrazující příčiny chybných detekcí PAC v rámci trénovacích dat. Všechny cykly zařazené do shluku PAC byly rozřazeny podle anotací z UBMI. Nejčastější patologií způsobující falešně pozitivní detekce v rámci shluku PAC je fibrilace síní, u níž je počet detekcí dokonce vyšší než počet pravdivě pozitivních detekcí tohoto shluku. Můžeme také vidět falešně pozitivní detekce způsobené záměnou PVC za PAC či přítomností šumu, což mohlo vést na falešné detekce vrcholu kmitu R a zkrácení RR intervalu. Do skupiny Ostatní spadají všechny zbylé falešně pozitivní cykly, u kterých nebyla zjišťována příčina.

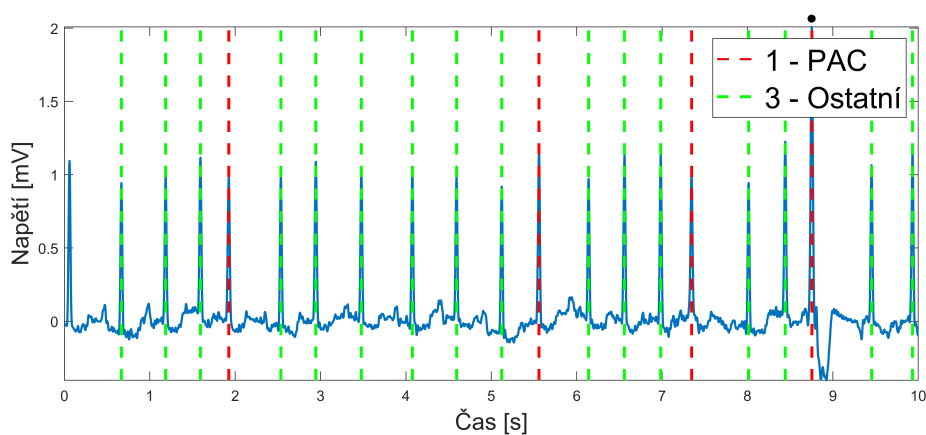


Obr. 7.2: Sloupcový graf zobrazující příčiny chybných detekcí PVC v rámci trénovacích dat. Všechny cykly zařazené do shluku PVC byly rozřazeny podle anotací z UBMI. Nejčastější patologií způsobující falešně pozitivní detekce v rámci shluku PVC je přítomnost šumu, která způsobuje zkreslení příznaků a následně chybné zařazení cyklu. I zde je určité procento falešně pozitivních detekcí způsobeno záměnou PAC za PVC. Nejmenší skupinu tvoří raménkové blokády (BBB) a AV blokády (AVB), které byly ve fázi návrhu algoritmu zařazeny do problematických kategorií z důvodu rozšířeného QRS komplexu a jejich podobnosti s PVC. Navrženým algoritmem se však povedlo tyto problematické patologie obejít.

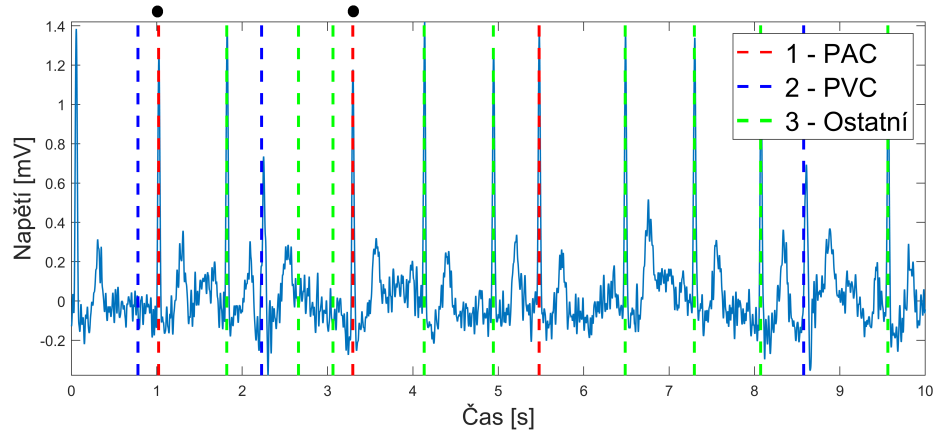
Jak již bylo zmíněno výše, snaha byla navrhnout algoritmus odlišný od dnešních moderních metod založených nejčastěji na umělé inteligenci. Právě z důvodu popularity umělé inteligence a také odlišnosti výše navrženého přístupu nebylo nalezeno velké množství podobných prací. Samozřejmě nalezneme řadu studií založených na shlukovacích algoritmech, avšak shlukování probíhá na základě podobnosti příznaků. Nalezeny byly také práce využívající korelaci, avšak většinou pro získání příznaků.

Z výše zmíněných důvodů není možné provést korektní srovnání výsledků s jinými studiemi. Avšak byly nalezeny studie založené na přístupu připomínající výše popsany. Přístupy v rámci těchto studií jsou založeny na korelaci jednotlivých cyklů s předem vyrobenými šablonami. Při získání vysoké hodnoty korelačního koeficientu jsou tyto cykly zařazeny do kategorie dané touto šablonou. Můžeme tedy říct, že jde o postupné tvoření shluků. Nalezené metody se liší v počtu šablon, se kterými jsou cykly korelovány. Studie dosahují velmi dobrých výsledků, které se v rámci všech prací většinou pohybují nad 95 % (TPR a FPR). Velkou výhodou těchto metod je možnost detekce v reálném čase, kdy se jedná o méně výpočetně náročné algoritmy, než v případě umělé inteligence. [18]-[20]

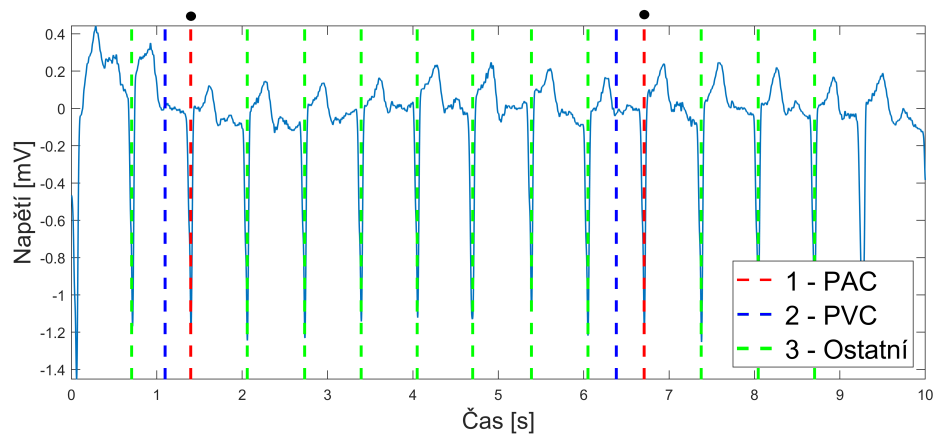
Nakonec jsou ještě uvedeny výsledky algoritmu u některých vybraných záznamů. Už dříve bylo zmíněno, že falešně pozitivní detekce mohou způsobit vzájemné záměny PAC a PVC (viz Obrázek 7.3). Důvodem mohou být stejné charakteristiky RR intervalů, či malá odlišnost PVC od ostatních cyklů. V databázi se nacházejí také záznamy obsahující detekce i v místech, kde se nenachází R kmit. I tato skutečnost může vést k FP výsledkům detekce PAC (viz Obrázek 7.4) i PVC (viz Obrázek 7.5). Jak je vidět na Obrázku 7.6 může se také stát, že již při shlukování je korelační koeficient kratšího segmentu tak vysoký, že i při odečtení hodnoty 0.15 bude stále nebývat nadprahových hodnot. Nakonec jsou uvedené záznamy, kde byla provedena úspěšná klasifikace PVC i PAC (viz Obrázky 7.7 a 7.8).



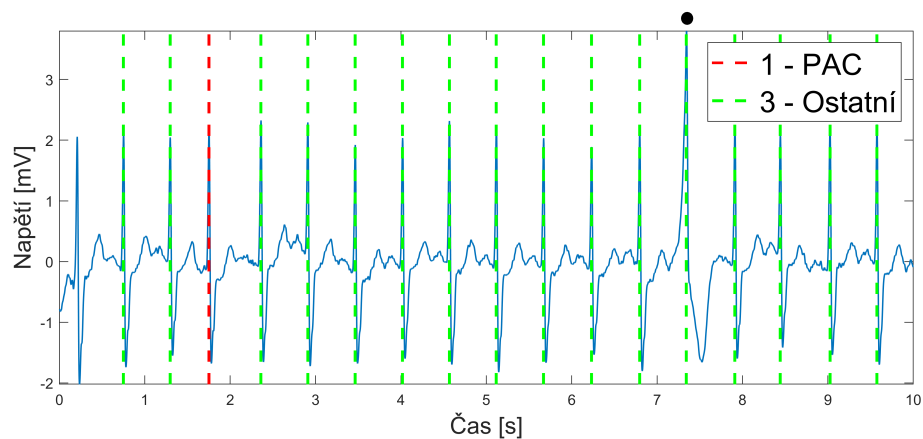
Obr. 7.3: Ukázka falešně pozitivní detekce PAC (označeno tečkou) kvůli záměně PVC za PAC. PVC nebyla v rámci shlukování rozpoznána a díky jejím charakteristickým změnám RR intervalů byla zařazena do PAC.



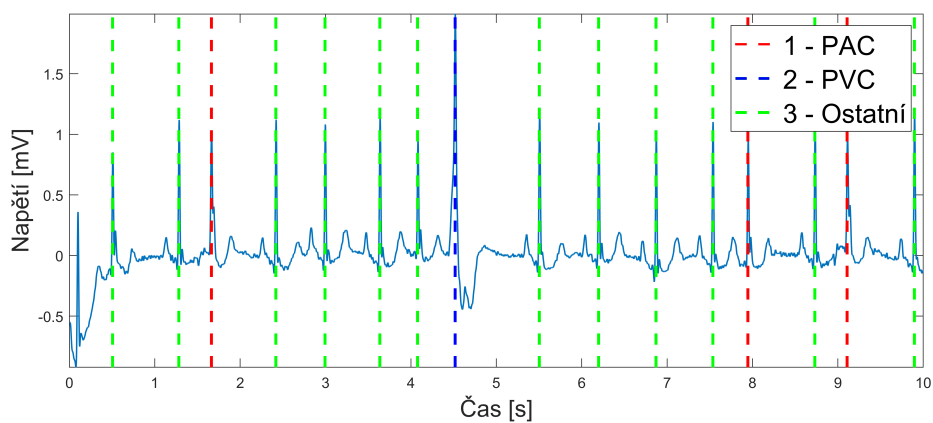
Obr. 7.4: Ukázka falešně pozitivní detekce (označeno tečkou) PAC kvůli detekci R kmitu v místě, kde se nenachází. Tyto detekce způsobily zkrácení RR intervalů a algoritmus následně vyhodnotil cykly za těmito detekcemi jako PAC. Jedna z chybných detekcí R kmitu (první detekovaná) byla vyhodnocena jako PVC nejspíše díky své odlišnosti do ostatních cyklů.



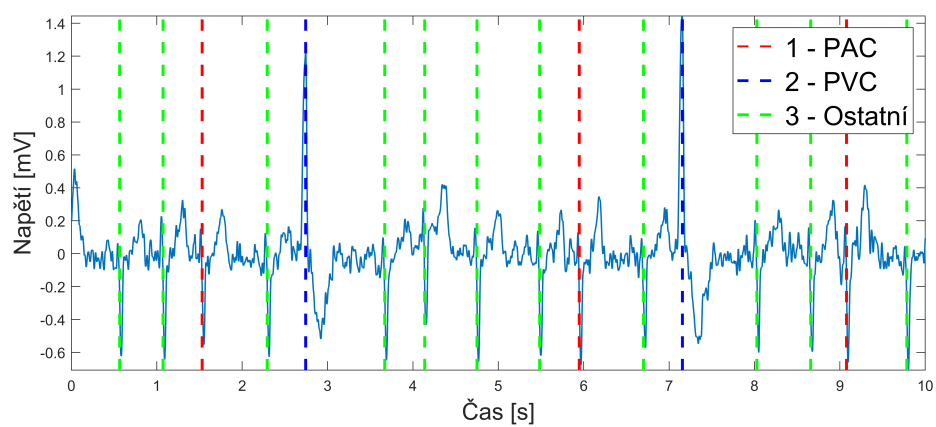
Obr. 7.5: Ukázka falešně pozitivní detekce PVC kvůli detekci R kmitů v místě, kde se nenachází. Navíc došlo ke zkrácení RR intervalů cyklů za těmito detekcemi a následně k jejich zařazení do shluku PAC (označeno tečkou).



Obr. 7.6: Ukázka chybné funkce algoritmu. V tomto prípade PVC nebyla detekována (označena tečkou) již v rámci shlukování na základě korelace. Korelace kratších segmentů dosahovala tak vysokých hodnot, že i po odečtení hodnoty 0.15 korelační koeficient nabýval stále nadprahových hodnot a zůstal v nejpočetnějším shluku, který se již dále nezpracovává.



Obr. 7.7: Ukázka správné funkce klasifikátoru, kdy můžeme vidět správně provedenou klasifikaci jednotlivých cyklů.



Obr. 7.8: Ukázka správné funkce klasifikátoru, kdy můžeme vidět správné zařazení jednotlivých cyklů do odpovídajících shluků.

Závěr

Cílem práce bylo navrhnout a následně realizovat algoritmus ve vývojovém prostředí MATLAB na základě metod shlukové analýzy, který měl provádět třídění jednotlivých EKG cyklů. Práce se měla zaměřit na detekci síňových a komorových extrasystol. Dílčím cílem práce bylo co nejvíce přiblížit rozhodování algoritmu postupu zdravotnického personálu při vyhodnocování záznamu EKG. Předpokládaný výsledek měl být v podobě číselného vektoru, kde by byly QRS komplexy zastoupeny čísly, která označují shluk, do něhož byl cyklus zařazen. V rámci bakalářské práce byl cíl naplněn. Algoritmus byl navržen tak, aby v každém záznamu provedl roztrídění cyklů do tří shluků a výsledek podal ve formě číselného vektoru, kde je každý cyklus zařazen pomocí konkrétního čísla (1-PAC, 2-PVC, 3-Ostatní). Důležité je také zmínit, že každý krok tohoto programu lze logicky vysvětlit a odůvodnit jeho provedení a proto bych ráda zmínila i splnění dílčího cíle o přiblížení se rozhodování zdravotnického personálu.

Jak již bylo zmíněno výsledek praktické části je v podobě číselného vektoru, kde každé číslo reprezentuje jeden komplex QRS a určuje jeho zařazení do konkrétního shluku. Ve zvoleném algoritmu je finálně možný vznik maximálně tří shluků (v rámci shlukování na základě korelace bylo množství shluků neomezené) a tím pádem rozdělení komplexů do tří skupin. Ideálně tyto tři skupiny představují shluk pro PAC, PVC a pro skupinu Ostatní.

Z kapitoly 7 můžeme usuzovat, že navržený algoritmus nedosahuje příliš kvalitních výsledků. V případě PAC to je způsobeno hlavně přítomností fibrilace síní, která způsobuje převážnou část falešně pozitivních detekcí. Silněji zašuměná data zůstávají stále problematickými a nevyřešeny zůstávají i problémy, kdy se PVC dostává do shluku pro PAC (díky krátkým RR intervalům), nebo detekce PAC v případě zařazení mimo nejpočetnější shluky. Možným nedostatkem navrženého algoritmu může být také nedostatečné rozměření signálu. Pokud by se provedlo detailnější rozměření, byla by možnost použití více příznaků, které by mohly přispět ke kvalitnějším výsledkům.

Navržený algoritmus má také své výhody. Jak již bylo zmíněno každý jeho krok je logicky vysvětlitelný a do jisté míry kopíruje rozhodování zdravotnického personálu. Při odstranění dosavadních nedostatků by mohl dosahovat dostatečně kvalitních výsledků, aby mohl být nápomocný v kardiologické praxi při hodnocení EKG záznamů. Pro lékaře by mohl být přístupnější už jen z důvodů hodnocení záznamů EKG podle pravidel, pro které byla inspirace právě v kardiologické praxi.

Literatura

- [1] HAMPTON, John R. *EKG stručně, jasně, přehledně* 7. Praha: Grada, 2013. ISBN 978-802-4742-465.
- [2] ZEMAN, Karel. *Poruchy srdečního rytmu v intenzivní péči*. Brno: Institut pro další vzdělávání pracovníků ve zdravotnictví, 1996. ISBN 80-701-3222-1.
- [3] ŠTEJFLA, DRSC., Prof. MUDr. Miloš, Doc. MUDr. Josef ŠUMBERA, CSC. a Prof. MUDr. Pavel BRAVENÝ, CSC. *Základy elektrokardiografie*. Brno: Masarykova univerzita v Brně, lékařská fakulta, 1991. ISBN 8021002654.
- [4] TECHmED: EKG & Arytmológia [online]. Bratislava, 2018 [cit. 2020-04-02]. Dostupné z: <https://www.techmed.sk/ekg-a-arytmologia-kniha/>
- [5] LUZ, Eduardo José da S., William Robson SCHWARTZ, Guillermo CÁMARA-CHÁVEZ a David MENOTTI. *ECG-based heartbeat classification for arrhythmia detection: A survey*. Computer Methods and Programs in Biomedicine [online]. 2016, 127, 144-164 [cit. 2019-12-13]. DOI: 10.1016/j.cmpb.2015.12.008. ISSN 01692607.
- [6] MARŠÁNOVÁ, Lucie, RONZHINA Marina, SMÍŠEK Radovan, VÍTEK Martin, NĚMCOVÁ Andrea, SMITAL Lukáš a NOVÁKOVÁ Marie. ECG features and methods for automatic classification of ventricular premature and ischemic heartbeats: A comprehensive experimental study. Scientific Reports. 2017, 7(1) [cit. 2020-05-23]. DOI: 10.1038/s41598-017-10942-6. ISSN 2045-2322.
- [7] SMÍŠEK, Radovan, HEJČ Jakub, RONZHINA Marina, NĚMCOVÁ Andrea, MARŠÁNOVÁ Lucie, KOLÁŘOVÁ Jana, SMITAL Lukáš a VÍTEK Martin. *Multi-stage SVM approach for cardiac arrhythmias detection in short single-lead ECG recorded by a wearable device*. Physiological Measurement [online]. 2018, 39(9) [cit. 2020-04-02]. DOI: 10.1088/1361-6579/aad9e7. ISSN 1361-6579.
- [8] LUZ, Eduardo José da S., William Robson SCHWARTZ, Guillermo CÁMARA-CHÁVEZ a David MENOTTI. *ECG-based heartbeat classification for arrhythmia detection: A survey*. Computer Methods and Programs in Biomedicine . 2016, 127, 144-164 [cit. 2019-12-13]. DOI: 10.1016/j.cmpb.2015.12.008. ISSN 01692607.

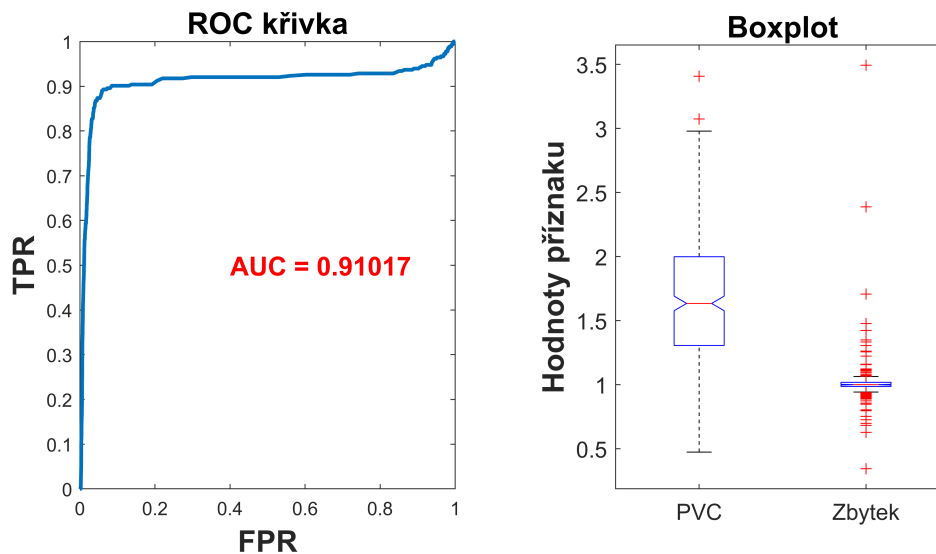
- [9] ESCALONA-MORAN, Miguel Angel, Miguel C. SORIANO, Ingo FISCHER a Claudio R. MIRASSO. *Electrocardiogram Classification Using Reservoir Computing With Logistic Regression*. IEEE Journal of Biomedical and Health Informatics [online]. 2015, 19(3), 892-898 [cit. 2020-04-14]. DOI: 10.1109/J-BHI.2014.2332001. ISSN 2168-2194.
- [10] LI, Qichen, Chengyu LIU, Qiao LI, Supreeth P SHASHIKUMAR, Shamim NEMATI, Zichao SHEN a Gari D CLIFFORD. *Ventricular ectopic beat detection using a wavelet transform and a convolutional neural network*. Physiological Measurement [online]. 2019, 40(5) [cit. 2020-04-14]. DOI: 10.1088/1361-6579/ab17f0. ISSN 1361-6579.
- [11] RONZHINA, Marina. *Umělá inteligence v medicíně: Shluková analýza* [přednášky]. Brno: Fakulta elektrotechniky a komunikačních technologií VUT. 1. října 2019.
- [12] ŘEZANKOVÁ, Hana, Dušan HÚSEK a Václav SNÁŠEL. *Shluková analýza dat*. 2., rozš. vyd. Praha: Professional Publishing, 2009. ISBN 978-808-6946-818.
- [13] VÍTEK, Martin. *Číslíkové zpracování signálů a obrazů - Korelační analýza signálů* [přednáška]. Brno: Fakulta elektrotechniky a komunikačních technologií VUT. 1.listopadu 2018.
- [14] KOZUMPLÍK, Jiří. *Analýza biologických signálů*. prezentace k přednáškám předmětu AABS
- [15] RONZHINA, Marina. *Klasifikace mikrosprávku analýzou EEG* [online]. Brno, 2009 [cit. 2020-04-06]. Dostupné z: <http://hdl.handle.net/11012/12044>. Diplomová práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav biomedicínského inženýrství. Vedoucí práce Vratislav Čmiel.
- [16] JURA, Pavel. *Základy fuzzy logiky pro řízení a modelování*. Brno: Vysoké učení technické v Brně, Nakladatelství VUTiUM, 2003. ISBN 80-214-2261-0.
- [17] RONZHINA, Marina. *Umělá inteligence v medicíně: Fuzzy množiny, Fuzzy logika, Fuzzy interferenční systémy* [přednášky]. Brno: Fakulta elektrotechniky a komunikačních technologií VUT. 26. listopadu 2019.
- [18] LIU, Chengyu, Peng LI, Yatao ZHANG, Yuan ZHANG, Changchun LIU a Shoushui WEI. *A Construction Method of Personalized ECG Template and Its Application in Premature Ventricular Contraction Recognition for ECG*

- Mobile Phones*. World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, 2013, 8(1), 585-588 [cit. 2020-05-23]. IFMBE Proceedings. DOI: 10.1007/978-3-642-29305-4_153. ISBN 978-3-642-29304-7. ISSN 1863-1703.
- [19] CHIU, CHUANG-CHIEN, TONG-HONG LIN, BEN-YI LIAU, Yuan ZHANG, Changchun LIU a Shoushui WEI. *USING CORRELATION COEFFICIENT IN ECG WAVEFORM FOR ARRHYTHMIA DETECTION*. Biomedical Engineering: Applications, Basis and Communications [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, 2013, 17(03), 147-152 [cit. 2020-05-23]. IFMBE Proceedings. DOI: 10.4015/S1016237205000238. ISBN 978-3-642-29304-7. ISSN 1016-2372.
- [20] LI, Peng, Chengyu LIU, Xinpei WANG, Dingchang ZHENG, Yuanyang LI a Changchun LIU. *A low-complexity data-adaptive approach for premature ventricular contraction recognition*. Signal, Image and Video Processing [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, 2013, 8(1), 111-120 [cit. 2020-05-23]. IFMBE Proceedings. DOI: 10.1007/s11760-013-0478-6. ISBN 978-3-642-29304-7. ISSN 1863-1703.
- [21] KAPOUNOVÁ, Jana a Pavel KAPOUN. *Bakalářská a diplomová práce: od zadání po obhajobu*. 1. Praha: Grada, 2017. ISBN 978-80-271-0079-8.

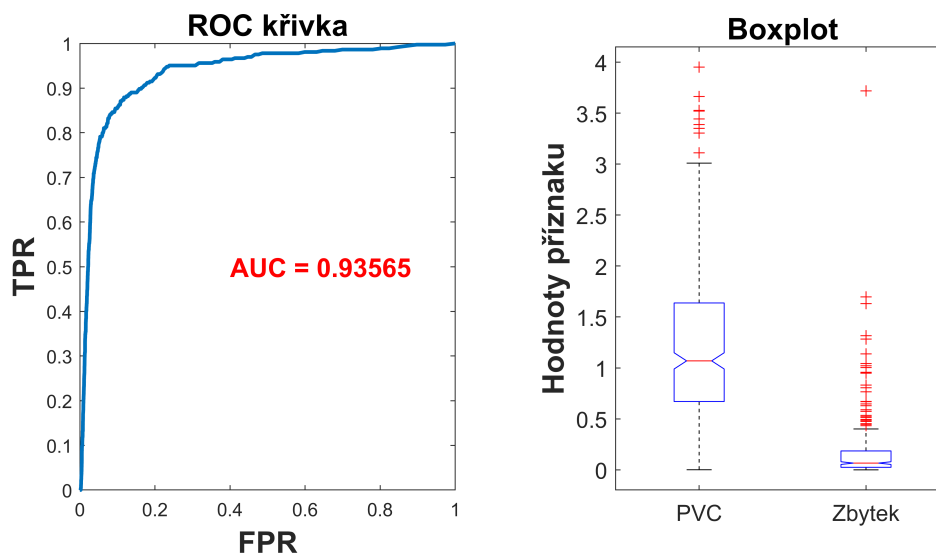
Seznam symbolů, veličin a zkratek

EKG	elektrokardiogram
AV	atrioventrikulární
SA	sinoatriální
PAC	síňová extrasystola
PVC	komorová extrasystola
AVB	atrioventrikulární blokáda
BBB	raménková blokáda
UBMI	ústav biomedicínského inženýrství
FPR	false positive rate - falešně pozitivní detekce
TPR	true positive rate - pravdivě pozitivní detekce
PPV	positive predictive value - pozitivně prediktivní hodnota
ACC	accuracy - přesnost

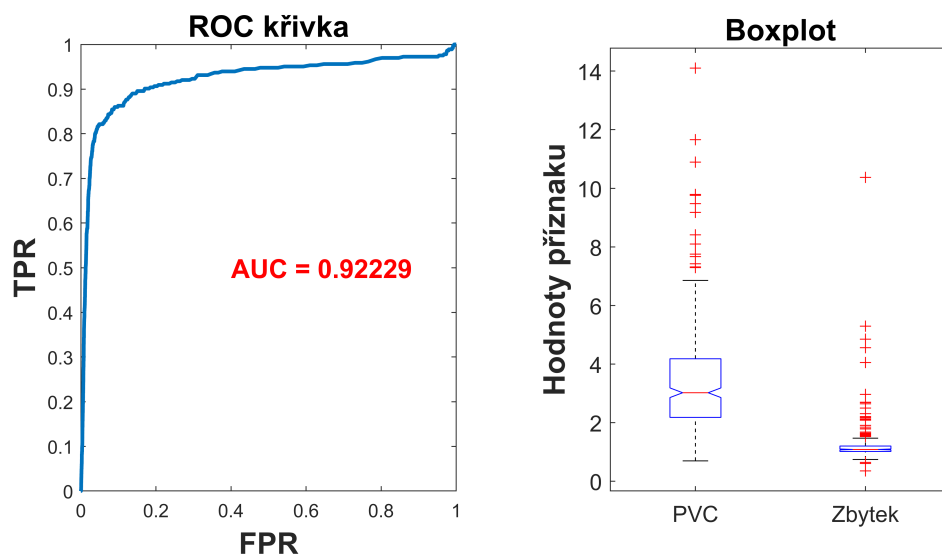
A ROC křivky a boxploty vybraných příznaků



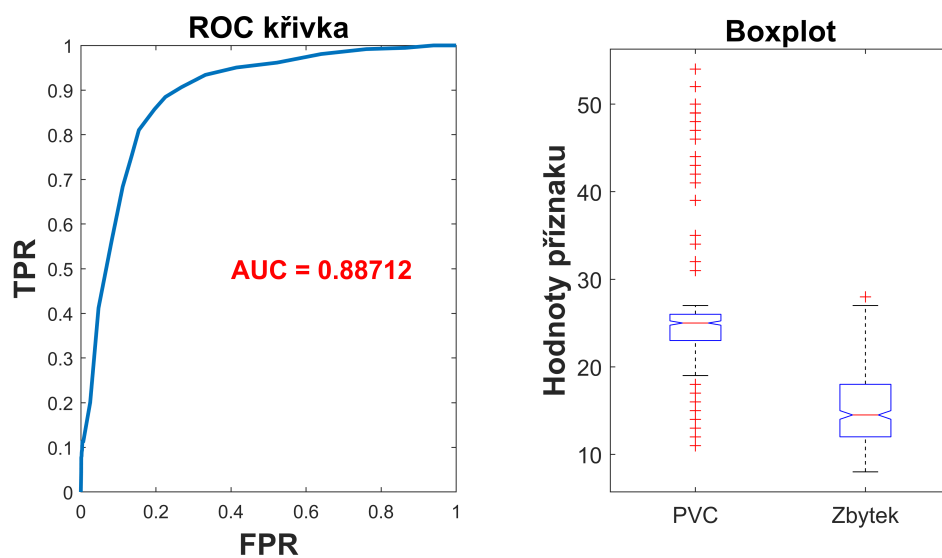
Obr. A.1: Ukázka ROC křivky a boxplotu pro příznak č.1: podíl RR intervalů



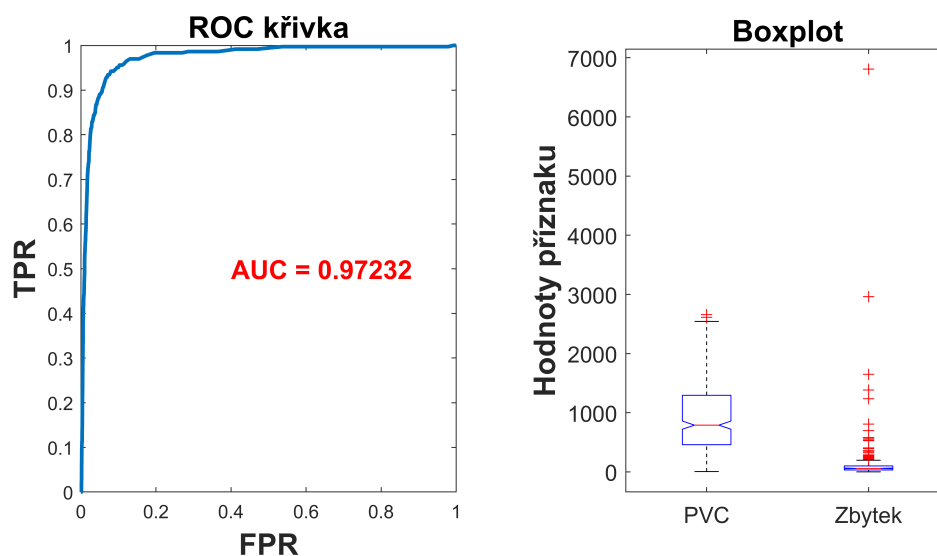
Obr. A.2: Ukázka ROC křivky a boxplotu pro příznak č.2: rozdíl tvarových faktorů cyklu a reference



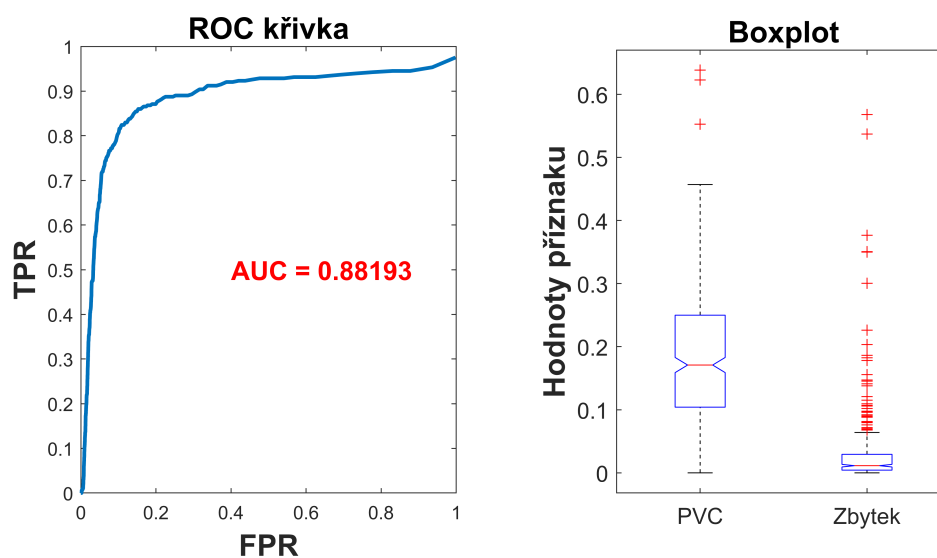
Obr. A.3: Ukázka ROC křivky a boxplotu pro příznak č.3: rozdíl obsahu ploch pod křivkou cyklu a reference



Obr. A.4: Ukázka ROC křivky a boxplotu pro příznak č.4: počet vzorků mezi extrémními cykly



Obr. A.5: Ukázka ROC křivky a boxplotu pro příznak č.5: podíl rozdílu napětových špiček maxima cyklu a reference ku rozdílu napětových špiček minima cyklu a reference



Obr. A.6: Ukázka ROC křivky a boxplotu pro příznak č.7: rozdíl intervalu RR před QRS komplexem a průměrnou hodnotou RR intervalu v rámci záznamu